Using Non-Linear/Chaotic Dynamics for Interest Rate Determination

Julian H. Tice

Thesis submitted for the Degree of PhD Industrial and Business Studies

September 1998

Supervised by Dr. N.J. Webber

TABLE OF CONTENTS

List of Figures

List of Tables

List of Algorithms

Acknowledgements

I would like to thank various people that helped me throughout my PhD. The most credit goes to Dr. N.J. Webber, my supervisor, for his constant support and guidance. I should also like to thank all the members of the department who have at one time or another contributed ideas and suggestions. The critical appraisal I have received from presentation and attendance at seminars and conferences has also been inspirational. Additional thanks go to Simon Babbs and Ben Nowman for their assistance in the provision of data and guidance regarding their Kalman filtering methods.

As to the people that are dear to me, I would especially like to thank my family and my partner, Cláudia S. Sarrico, for their continual encouragement and support.

This research was made possible by the financial support of the ESRC, award no. R00429334359, and the supplementary support of Warwick Business School.

Declaration

During my PhD studies the following research paper was published :

Tice J.H. and N.J Webber (1997) A Non-linear model of the term structure of interest rates, *Mathematical Finance*, **7**(**2**): 177-209

Summary

A new class of interest rate models is proposed where the main driving terms for the large scale dynamics of the system are deterministic. As an example, an economically motivated two factor model of the term structure is presented that generalises existing stochastic mean term structure models.

By allowing a certain parameter to acquire dynamical behaviour the model is extended to three factors. It is shown, that in a deterministic version, the model is equivalent to the Lorenz system of differential equations. With reasonable parameter values the model exhibits chaotic behaviour. It successfully emulates certain properties of interest rates including regime switching and behaviour of a business cycle nature. Pricing and term structure issues are discussed. Standard PCA techniques used to estimate HJM type models are observed to be equivalent to dimensional estimates commonly applied to 'spatial data' in non-linear systems analysis. An empirical investigation uncovers surprising structure consistent with the existence of a low dimensional attractor. Issues of control of the chaotic system with reference to the underlying economic model are discussed.

A heuristic approach is made to estimating the three factor model. Exploiting properties of the term structure, the existence of noise, and the geometry of the system allows a variety of methods for uncovering the parameters of the model. A better approach is found in the application of the Kalman filter to the estimation problem. Lack of explicit solutions motivates an investigation into the use of approximated forms for the term structure. The traditional Kalman filter is seen to be unstable when applied to the chaotic three factor model. A stable variant, from the class known as 'square-root' filters, is adopted. A new method is created for finding the analytical derivatives of the log-likelihood function such that it is consistent with the 'square-root' filter. Estimates for the empirical estimation of the models developed earlier in the thesis are given.

It is concluded that there is much scope for expanding the literature within the new class of models proposed. The particular three factor model developed has been shown to have realistic properties and be amenable to bond and contingent claim pricing. The chaotic nature of the model, underpinned by an economic derivation, opens up new methods for authorities to control/stabilise the economy. An analysis of the underlying dynamical structure of UK money market rates is consistent with a low dimensional deterministic driving force. Heuristic methods employed to estimate the parameters of the model allow for an insight into exploiting the geometry of the system. Application of the Kalman filter to estimation of non-linear models is found to be problematic due to the linearisations/approximations that are necessary.

An outline of areas for future research is given, providing ideas for extending the economic formulation, further investigating control of chaotic interest rate systems, testing empirical data for evidence of chaotic invariants and methods for quantifying and improving the Kalman filtering procedures for better handling non-linear models.

1. INTRODUCTION

The use of complex dynamics to describe economic and financial processes is not a new subject, despite the relative paucity of literature in the area. It has long been known that the types of behaviour non-linear systems are capable of possessing are well suited to describing financial and economic processes. Despite this, much research concentrates on systems with limited dynamic behaviour. To an extent, this has been motivated by the need to apply theoretical frameworks to real world data. Describing non-linear phenomena in empirical data is still a contentious area and the analytical tractability of systems which provide closed form solutions is very appealing.

Models of the term structure of interest rates currently used in finance fall into several categories. In most the motivation is to fit some feature or other of the behaviour of the term structure. For instance, some models, such as Vasicek (1977) and Cox, Ingersoll and Ross (1985a) (1985b), focus on describing the dynamics of the short rate. Others, such as Heath, Jarrow and Morton (1992), attempt to match the shape and the dynamics of the entire term structure. Yet another category, such as Babbs and Webber (1994), and Balduzzi, Bertola and Foresi (1993) attempt to model the rate setting behaviour of the monetary authorities.

It is possible to extend models of the short rate to capture some of the features of models that fit the current term structure. For instance, extended Vasicek models permit time varying behaviour of the mean level to which the short rate reverts. In the Hull and White (1990a) model, and its extensions (for example Babbs (1993)), the reversion level (and some other parameters) are allowed to be functions of time. This enables the model to fit, amongst other things, an arbitrary current term structure.

A number of studies, including Chan, Karolyi, Longstaff and Sanders (1992), have investigated the empirical behaviour of the short rate. One of the most evident features is level dependent volatility, which, in the models estimated by CKLS is captured through the volatility terms having r dependence. However, level dependent volatility can also be produced through non-linear models with recourse to complex volatility forms. It is fiurther found that there appears to be only weak evidence for the existence of a long run level of reversion. Since the short rate seems to be a stationary process, this suggests that the short rate reverts to a short run mean that may be changing through time.

Some models choose a process for the short rate which reflects the fact that the mean of the process may itself be time varying. Hull and White (1994a) (1994b), Sørensen (1994) and Chen (1996) are examples where the mean is itself a mean-reverting process. Introducing a stochastic mean allows the range of term structures that can be fitted by the model to be considerably expanded. Other dynamical structures for the mean have also been formulated. The two-factor model of Longstaff and Schwartz (1992) has a stochastic short rate and a stochastic short rate volatility, v. The short rate mean-reverts to an affine function of v, and v mean-reverts to an affine function of the short rate. Another example of a model with a more complex process for the mean is that of Brennan and Schwartz (1979). In addition to the short rate, Brennan and Schwartz allowed a long rate, the yield to maturity on a perpetual coupon bond, to be a second stochastic factor. The short rate mean-reverts to a linear function of this long rate[1].

Although the extended Vasicek and Cox, Ingersoll and Ross models, and the stochastic mean models, have had some success in modelling the behaviour of interest rates, these have been developed more for their ability to fit dynamic and static features of the term structure rather than their ability to account for interest rates as arising from fundamental economic processes. Notable exceptions are the models of Cox, Ingersoll and Ross, and Longstaff and Schwartz which are explicitly derived from a general equilibrium model of an economy. This framework not only guarantees no-arbitrage,

[1] Hogan (1993) subsequently showed that the complete system of Brennan and Schwartz was unstable.

but also provides an intuitive justification for the form of the model. While, as Duffie and Kan (1994, 1996) point out, almost any model, with suitable regularity conditions, may be considered to have arisen out of a general equilibrium framework, nevertheless there has been relatively little written about the relationships between these models and concepts from economics. Furthermore, little attention has been paid to displacements from general equilibrium in these models. The models presented in this thesis, through not set in a general equilibrium framework, attempt to address some of these issues.

1.1 Should the short term interest rate drift be considered non-linear?

This research focuses on dynamics; the study of change and the forces generating it. Simple dynamics are exemplified by stationary states, periodic cycles or balanced growth or decline. However, such simple dynamics are not generally reflective of empirical financial series. Non-linear dynamics are capable of producing change that is not balanced, encompassing such features as nonperiodic fluctuations, overlapping waves, switching regimes and structural change. The irregular nature of financial data seems often better described by such an approach.

An intrinsic feature of this approach to characterising financial series is that it does not rely on external shocks to produce the random nature observed in the data. The stochastic behaviour is explained endogenously within the model. Whilst it would be foolish to posit that all of the random nature of such financial series can be accounted for with such methods, it can provide a plausible explanation for at least a significant part of it. In financial markets we often observe periods of regularity interspersed with abnormal events. As an example of this, Peters (1994) notes the way in which investment horizons change with the stability of the market. If a questionable event hits the market, such that the long term earnings power of firms cannot be agreed upon, the market stability collapses. In this state, short term information may cause price movements that are larger than would be observed with a stable market. As the market regains liquidity with investment horizons widening, investors return to market fundamentals and economic factors for information, restoring market stability. Peters refers to such a system as one with global structure but local randomness, with the system switching between stable and unstable regimes. As well as regime switching, non-linearity is naturally able to account for a variety of statistical anomalies in interest rate data that do not fit the linear specification. These may include time varying risk premia, level dependent volatility and time varying persistence of shocks.

Intuitively, it is highly plausible for a complex financial process such as the short term interest rate to be described by a non-linear functional form. However, much of the literature is pre-occupied with term structure models comprising linear drift functions. A priori, it appears somewhat anomalous to impose such a restriction on a process which is created by many interacting underlying economic forces. From this standpoint, the onus of justification should be on those wishing to assume a linear functional form. Much of the reason for the restrictive assumptions reduces to a question of analytical tractability. Many linear forms for the short rate process exhibit closed form solutions. Only in particular circumstances are explicit solutions available for non-linear models. The intractable analytical properties of the class of non-linear models makes them undesirable for application and estimation. Processes are chosen for their ability to model the statistical and probabilistic features of the empirical rates. Economic considerations are given a minor role in the derivation of the functional form and for the most part the resultant process parameters have no economic significance attributable to them.

Recently, some investigations have been conducted to ascertain whether the drift function of the short rate contains non-linearities. Stanton (1997) and Aït Sahalia (1996) have proposed nonparametric estimators of the drift and diffusion functions. They respectively investigate US Treasury bill yields and Eurodollar rates. Both authors find that the estimated drift function is highly non-linear. In a re-evaluation of these results Chapman and Pearson (1998) apply these nonparametric estimators to simulated sample paths of a square root diffusion process (with linear drift) finding that the estimated drift functions display non-linearities of the type reported by Stanton and Aït Sahalia. Chapman and Pearson conclude that it is difficult to use the estimators of Stanton and Aït Sahalia to produce reliable inferences concerning the presence of non-linearity in the short rate drift. GMM estimation applied to the data sets finds little evidence of non-linearity for the Treasury bill yields, whilst evidence of non-linearity is found for the Eurodollar rate but of the opposite sign to that found by Aït Sahalia. The overall conclusion drawn is that current estimation techniques are not robust enough to give irrefutable evidence regarding the linearity or non-linearity of the short rate drift.

An earlier study by Mizrach (1996) proposes a non-linear term structure model based on a generalisation of the multi-factor approach of Langetieg (1980). The model is tested nonparametrically and compared against its linear counterpart. He finds that in sample the non-linear model does not perform significantly better, but out of sample sees a significant improvement in forecast performance. Pfann, Schotmann and Tschernig (1996) take the approach of modelling US Treasury bill data using an autoregressive model with different regimes. This formulation has the advantage of modelling empirical properties such as time varying persistence of shocks and level dependent volatility. They find the presence of different regimes operating at low and high levels of the short rate. This finding is broadly consistent with that of Stanton (1997).

Overall, the approaches taken to investigating the applicability of non-linear models to the term structure of interest rates have produced mixed results. Undoubtedly, this is at least in part due to the complexity of the analytical techniques involved. Not enough is known about the finite sample properties of many of the estimators that are applied. How much the results are reflective of the estimation procedure is often difficult to tell.

1.2 Contribution and outline

There are four substantive areas of contribution from this work. Firstly, it is desired to illustrate how an interest rate model, derived from economic relationships, may exhibit complex dynamics. Moreover, the form of the dynamics generated will be seen to be comparable with those generated by the real world process. Thus, economic meaning is ascribed to the large scale fluctuations of the model, with the volatility terms representing true "noise". It will be demonstrated that a class of two factor financial models of the short rate may be thought of as set within the economic framework discussed. This class is quite broad, including the time dependent mean models of Hull and White (1990a) (1994a) (1994b), Sørensen (1994) and others, although it excludes models that incorporate stochastic volatility.

Secondly, a particular example of a model is described in which the short rate exhibits chaotic behaviour, switching from regimes of high rates to regimes of low rates seemingly at random. In this model, as in the class as a whole, a stochastic term is interpreted as true noise and is not in itself responsible for large scale fluctuations in the short rate. In the particular model the main source of large scale variation in the short rate, causing swings back and forth between high and low rates, is due to the deterministic term and not the stochastic term.[2] It is believed that this is the first naturally derived model of the short rate exhibiting chaotic behaviour.

The third contribution of this thesis is the observation that techniques used to investigate 'spatial data' for non-linear features are already employed in interest rate modelling, albeit without an apparent awareness of their significance. This implies that various mathematical methods, not previously applied to interest rate modelling, may

[2] The usual nomenclature of 'drift term' seems inappropriate for this model in that the drift term is a driving term that causes, in some sense, the main variation in the model.

now be deliberately employed. This could lead to new insights into interest rate dynamics.

Methods for estimating the models proposed are investigated, providing a fourth contribution. The loss of analytical tractability from the complex nature of the models prompts a variety of investigative techniques. The use of filtering techniques is investigated. In particular, interest focuses on the lack of explicit solutions and absence of traditional model stability that comes with complex dynamics. A new method is developed for providing consistent analytical derivatives of the log-likelihood function, when the dynamical system is traditionally unstable.

The plan of this thesis is as follows: In chapter two, interest rate dynamics in a standard economic framework are examined, leading to the derivation of a two-factor model of short rate dynamics. It is possible to indicate the effect of monetary and fiscal economic policies upon the behaviour of interest rates and to relate this to time varying behaviour of the mean. In chapter three a certain parameter, representing the strength of influence between the current level of the short rate and the future state of the economy, is allowed to acquire dynamics. A deterministic version of this three factor model is demonstrated to be equivalent to the Lorenz system of differential equations, and so may exhibit chaotic behaviour with certain ranges of parameter values. It is established that economically plausible parameters do indeed produce chaotic behaviour. Estimation and pricing issues are discussed, showing how standard non-linear techniques are related to methods used to calibrate Heath, Jarrow and Morton type models. Chapter four investigates how an estimation procedure may proceed. Knowledge of the geometry of the particular system, and techniques from investigating chaotic systems are applied. The presence of noise is seen to be a beneficial element. A number of methods are discussed for estimating the individual parameters of the model.

Estimation is found to be hampered by the procedure of estimating parameters individually and subsequent recovery of the state variables. Chapter five seeks a compact and efficient estimation procedure in the form of the Kalman filter. Its ability to model the time evolution of the term structure allows large amounts of spatial data to contribute to the estimation procedure. A numerically superior class of filters, known as 'square-root' filters, is discussed and a particular variant is adopted. A new method for finding the derivatives of the log-likelihood function is developed, to be consistent with the 'square-root' filter employed. Application of the filter is made to the one and two factor generalised Vasicek processes of Babbs and Nowman (1997). The use of an approximation to the term structure is investigated and compared against its theoretical equivalent. The approximation is of interest as the exact solution to the bond pricing equation is not available for many models exhibiting complex dynamics. The particular two and three factor dynamic mean models developed in chapters two and three are estimated using the 'square-root' filter. An example of the failure of the conventional Kalman filter is given, with the 'square-root' filter performing consistently better.

It is concluded that dynamic mean models have the ability to produce dynamics qualitatively similar to those observed in a variety of financial processes. Furthermore, the class of dynamic mean models, through its economic underpinnings, allows for economic interpretation and control. Many techniques employed currently for investigating empirical term structures may be refined allowing for the uncovering of interest rate dynamics. Filtering techniques can be applied to dynamic mean models, although success may be limited. Linearisations and approximations necessary for many non-linear models may be detrimental to the estimation process. The length of the data series may also be a significant factor in determining reliable estimates, where the large scale dynamics form an intrinsic part of model design.

2. DYNAMIC MEAN INTEREST RATE MODELS AND THE IS-LM FRAMEWORK

2.1 Introduction

The objective here is to motivate an interest rate model on economic grounds which can display deterministic dynamics capable of qualitatively describing the large scale dynamics of the real world process. This provides two major advances over many interest rate models common in the literature. Firstly, many models for the short rate are motivated by their ability to model the statistical and probabilistic aspects of the real world process. Little or no attempt is made to give economic foundations to the model. Where foundations are derived with economic motivation, such as a general equilibrium framework, it is often the case that the particular form for the process chosen has little economic meaning attributable to it. Secondly, a large portion of the literature seeks to concentrate emphasis on the form of the volatility function for the short rate process. The drift term then has limited effect on the dynamics of the model, often serving the sole purpose of mean reversion. It seems desirable that the drift term, encompassing an economically derived form, should be able to describe the large scale dynamics of the short rate process. It is the use of such forms for the drift of the process, allowing complex dynamics, that are investigated here. This allows for the economically derived deterministic process to describe fluctuations such as the business cycle.

Section 2.2 describes a dynamic mean framework and how many of the existing models in the literature may be thought of as being couched within it. Section 2.3 presents a particular example of a dynamic mean model derived from the economic IS-LM framework. It is shown that the model generalises several popular interest rate models in the literature. Section 2.4 assess the possible dynamics of the model, showing that oscillatory dynamics capable of business cycle type behaviour are possible. The volatility function is seen to have a trivial role in the large scale dynamics of the process. Section 2.5 concludes.

2.2 A Dynamic Mean Framework

In this section a class of dynamic mean interest rate models is described. A brief outline of the IS-LM framework is given, showing how existing dynamic mean interest rate models may be interpreted within it.

Define a dynamic mean interest rate model to have the following form :

$$
\begin{aligned}
dr &= \alpha(x - r)dt + \sigma_r dz_r \\
dx &= \beta(\mu_x(t,r,Y) - x)dt + \sigma_x dz_x \\
dY &= \gamma(\mu_Y(t,r,Y) - Y)dt + \sigma_Y dz_Y
\end{aligned}
\qquad (2.2\text{-}1)
$$

where r is the short rate and x is the level to which the short rate reverts. Y is a vector process summarising the remainder of the dynamics in the model, via the function μ_Y. Table 2-1 shows how a number of existing models may be regarded as being of this form (although the vector Y is trivial in all of them).

TABLE 2-1: COMPARISON OF TERM STRUCTURE MODELS

The short rate process is $dr = \alpha(x - r)dt + \sigma_r dz_r$.

Model	Process for x.
Vasicek	$dx = \mu dt,\ \mu = 0$
Hull and White (90)	$dx = \mu(t)dt,\ \mu(t) = \dot{x}$
Chen, et al	$dx = \beta(\mu - x)dt + \sigma_x dz_x,\ \mu$ constant
Longstaff and Schwartz*‡	$dx = \beta(\mu(t,r) - x)dt$ + volatility terms
	$\mu(t,r) = a + br.$
Babbs and Webber†	$dx = \beta(\mu(t,r) - x)dt + \sigma_x dz_x$
	$\mu(t,r) = pr + (1-p)\mu,\ p$ constant

* After reparameterisation.
‡ The volatility structure is ignored here
† x serves a role analogous to the mean of the short rate

Vasicek may be regarded as a trivial example of a dynamic mean model. Longstaff and Schwartz is known as a stochastic mean model. Here the dynamics of its mean are of concern.

Note that in the definition of a dynamic mean interest rate model no assumptions are made concerning the form of the volatility functions σ_r, σ_x and σ_Y. The interest is in the structure of the dynamics of the mean of r, not in the volatility. Volatility functions shall be regarded as constants.

Dynamic mean interest rate models are well known for their ability to fit initial term structures, and with Gaussian volatilities may also yield explicit solutions for bond and bond option prices.

It is now demonstrated how examples of dynamic mean models may arise from the IS-LM framework. The IS-LM model is a long standing standard model in macro-economics first introduced by Hicks (1937). Modern treatments may be found in Dornbusch and Fischer (1994) or Blanchard and Fischer (1989). The framework is the basis for various studies in economics, such as the dynamics of economic systems (Dernburg and Dernburg (1969)) and mechanisms of exchange rate determination (Krugman and Miller (1992)). An economic setting for the IS-LM model is outlined before showing how it leads to a range of two factor dynamic mean term structure models.

Many relationships in economics are expressed as equalities 'in equilibrium'[1]. For instance, in equilibrium supply and demand for a good are equal, or national real income and expenditure are equal. If a disturbance takes the system out of a stable equilibrium then a restoring force tends to move the system back towards equilibrium. It can be supposed that an economic system may be described by a set of equilibrium equations of the form

$$dm_i = \alpha_i(\overline{m}_i - m_i)dt + \sum_{j=1}^{N} \sigma_{ij} dz_j, \quad i = 1 \ldots N \qquad (2.2\text{-}2)$$

where quantities m_i adjust to equilibrium values $\overline{m}_i$ at rates $\alpha_i > 0$ when subject to disturbances represented by the N independent standard Wiener processes z_j. The effect

[1] This is to be interpreted informally and does not necessarily mean, for instance, arising from a general equilibrium framework.

of the disturbance is then proportional to the time since it occurred; that is, it decays away in a monotonically decreasing fashion.

Further, suppose that there are N state variables, $X_1,\ldots,X_N$, one of which, X_1 say, is the short rate. The quantities $m_i = m_i(X_1,\ldots,X_N)$ and $\overline{m}_i = \overline{m}_i(X_1,\ldots,X_N)$ are functions of the state variables. From the dynamics (2.2-2), the dynamics of the state variables are inferred.[2] Here it is the case that

$$dX = \left(M_X^{-1}A\left(\overline{M} - M\right) - M_X^{-1}h\right)dt + M_X^{-1}\Sigma dz \tag{2.2-3}$$

where X, $\overline{M}$, M and dz are the vectors $\{X_i\}$, $\{\overline{m}_i\}$, $\{m_i\}$ and $\{dz_i\}$, A is the matrix $diag(\alpha_1,\ldots,\alpha_N)$, $\Sigma = \{\sigma_{ij}\}$, M_X is the matrix $\left\{\frac{\partial m_i}{\partial X_j}\right\}$, and $h = \{h_i\}$ is defined as

$$h_i = \frac{1}{2}\sum_{k,l,p=1}^{N} \frac{\partial^2 m_i}{\partial X_k \partial X_l}\sigma_{kp}^X \sigma_{lp}^X \tag{2.2-4}$$

where

$$\sigma^X = M_X^{-1}\Sigma$$

When M and $\overline{M}$ are affine functions of the state variables X,

$$\begin{aligned} M &= BX + \underline{b} \\ \overline{M} &= CX + \underline{c} \end{aligned} \tag{2.2-5}$$

the system *(2.2-3) simplifies to*

$$dX = B^{-1}A\left((C - B)X + (\underline{c} - \underline{b})\right)dt + B^{-1}\Sigma dz \tag{2.2-6}$$

2.3 The IS-LM Framework

A particular example of an affine model is the IS-LM framework. There are two state variables $X = (r, y)'$ where r is the real short rate and y is the real rate of national income. There are two equilibrium relationships. These equate in equilibrium (i)

[2] It is assumed that equations (2.2-1) possess non-degenerate solutions and may be inverted.

money supply, m_s, with money demand, m_d, and (ii) the real rate of national income, y, with the real rate of national expenditure, e. The equilibrium equations are:

$$\begin{aligned} dm_d &= \alpha_m (m_s - m_d)dt + \sigma_m dz_m \\ dy &= \alpha_y (e - y)dt + \sigma_y dz_y \end{aligned} \tag{2.3-1}$$

The way in which the economy moves to equilibrium depends on the speed of reaction of the money and goods market. It can be expected that interest rates will adjust every minute to regulate the demand in the money markets. However, prices in the goods market adjust only slowly such that equilibrium is restored much less quickly than the money markets following a disturbance. α_m is the rate of adjustment in the money market and is large. α_y is the rate of adjustment in the goods market and is small. m_d and e are functions of r and y and defined by

$$\begin{aligned} m_d &= ky - ur \\ e &= a - br + cy \end{aligned} \tag{2.3-2}$$

with $u, k, a, b, > 0$, $0 < c < 1$. Set $M = (m_d, y)'$, $\overline{M} = (m_s, e)'$, $m_s > 0$. In the IS-LM framework the coefficients B, C, $\underline{b}$ and $\underline{c}$ in equations 2.4 are

$$B = \begin{pmatrix} -u & k \\ 0 & 1 \end{pmatrix}, \; C = \begin{pmatrix} 0 & 0 \\ -b & c \end{pmatrix}, \; \underline{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \; \underline{c} = \begin{pmatrix} m_s \\ a \end{pmatrix},$$

and set

$$\Sigma = \begin{pmatrix} \sigma_m & 0 \\ 0 & \sigma_y \end{pmatrix}, \; A = \begin{pmatrix} \alpha_m & 0 \\ 0 & \alpha_y \end{pmatrix}.$$

Substituting into (2.2-6) it is found that

$$\begin{aligned} dr &= v\left(\frac{q}{v} + \frac{w}{v} y - r\right)dt + \sigma_r dz_r \\ dy &= \alpha_y (1-c)\left(\frac{a}{1-c} - \frac{b}{1-c} r - y\right)dt + \sigma_y dz_y \end{aligned} \tag{2.3-3}$$

where

$$q = a\alpha_y \frac{k}{u} - \alpha_m \frac{m_s}{u}$$

$$v = \alpha_m + \alpha_y b \frac{k}{u}$$

$$w = \alpha_m \frac{k}{u} - \alpha_y (1-c) \frac{k}{u}$$

$$\sigma_r^2 = \frac{1}{u^2}\left(\sigma_m^2 + k^2\sigma_y^2\right)$$

and the correlation between z_r and z_y is $\rho_{ry} = k\sigma_y / u\sigma_r$

Equations (2.3-3) give us the dynamics of y and r. Variations of this system have been investigated, for instance in Dernburg and Dernburg (1969) and Blanchard and Fischer (1989). The economic system presented here is a simplified view of the real world. One major omission in this model is the lack of inclusion of the supply side, allowing the price level to be determined endogenously. The implication here is that aggregate supply is infinitely elastic for a given price level, and that (2.3-3) describes a fix price Keynesian model. The underpinnings of the IS-LM model come in the way that the goods and money markets interact. The economy is conceptually divided into two macromarkets. The product market describes the flow of real output, while the money market describes the stocks of money, bonds and other financial assets held. Determination of expenditure in the product market influences demand for money in the money market. Similarly, the rate of interest determined in the money market plays an important role in influencing certain categories of expenditure in the product market. These interacting relationships are bound up in the system (2.3-3).

The parameters of the relationships (2.3-2) have economic interpretations. The demand for money is described by *speculative* and *transactions* demand. Investors can keep their liquid assets in either money or bonds. As the interest rate increases, the expectation is that investors will reduce their demand for money as they invest in bonds. Hence $dm_d/dr = -u,\ u > 0$, describes investors liquidity preference and represents speculative demand. Transactions demand represents the requirement for money to be held to pay for goods and services within the economy. As expenditure and income grow, transactions demand for money rises. Hence, $dm_d/dy = k,\ k > 0$. For the product market, expenditure is described by three components, these being autonomous expenditure, investment and consumption. The term a represents autonomous expenditure comprising fixed levels of government expenditure, investment and consumption within the economy. The term c represents the marginal propensity to consume out of income and, as such, takes on a value between 0 and 1. Investment in the economy is dependent on the level of the interest rate, such that a rise in the interest rate will cause a reduction in investment, as the cost of borrowing rises. Hence $dy/dr = -b,\ b > 0$.

The focus is upon the dynamics of r, so a change of variable will be performed to simplify the system. The change of variable also enables the re-expression of (2.3-3) in a form more familiar in the finance literature.

Set $x = \dfrac{q}{v} + \dfrac{w}{v}y$ and rewrite (2.3-3) in terms of x and r. The resultant expression in (2.3-4) will then be consistent with the dynamic mean framework form defined in (2.2-1). In essence the situation can be viewed as one in which the dynamics of r are determined by the behaviour of x. The facet that r reverts to the current value of x. implies that x will represent a prospective level of interest rates, a concept previously described in Babbs and Webber (1994, 1997). The process for x will then encapsulate the underlying economics. One obtains :

$$\begin{aligned} dr &= v(x-r)dt + \sigma_r dz_r \\ dx &= \alpha_y(1-c)\left(\frac{w}{v}\frac{a}{1-c} + \frac{q}{v} - \frac{w}{v}\frac{b}{1-c}r - x\right)dt + \sigma_x dz_x \end{aligned} \tag{2.3-4}$$

where $\sigma_x = \dfrac{w}{v}\sigma_y$, and $z_x = z_y$. To elucidate the underlying structure, set

$$p = -\frac{w}{v}\frac{b}{1-c},$$

$$\mu = \frac{a + \frac{q}{w}(1-c)}{b + \frac{v}{w}(1-c)}$$

and define $\alpha = v$ and $\beta = \alpha_y(1\text{-}c)$, then (2.3-4) becomes

$$\begin{aligned} dr &= \alpha(x-r)dt + \sigma_r dz_r \\ dx &= \beta\big(pr + (1-p)\mu - x\big)dt + \sigma_x dz_x \end{aligned} \qquad (2.3\text{-}5)$$

x mean reverts to a weighted sum of r and μ. With our sign assumptions, and since $\alpha_y < \alpha_m$, the weighting factor p is negative, and if c is close to 1, $|p|$ could be large.

2.4 Assessing the dynamics of the two factor model

One may consider the system (2.3-4) in the absence of noise. Setting σ_r and σ_x to zero the fixed points of (2.3-5) occur at r_0 and x_0, where

$$r_0 = x_0 = \mu, \qquad p \neq 1.$$

When $p = 1$ the line $x = r$ is an invariant submanifold. When $p = 0$, the equation for x becomes

$$dx = \beta(\mu - x)dt$$

This has solution

$$x(t) = x(0)e^{-\beta t} + \mu\left(1 - e^{-\beta t}\right),$$

thus (2.3-5) reduces to a system where r reverts to a time dependent mean. If μ is allowed to be an arbitrary integrable time dependent function, then the value to which r reverts,

$$x(t) = x(0)e^{-\beta t} + e^{-\beta t}\beta\int_0^t e^{-\beta t}\mu(t)dt,$$

can be an arbitrary differentiable function of time. This is reminiscent of the time dependent mean model of Hull and White (1990a).

The system described by (2.3-5), where p is an arbitrary constant, generalises the drift functions assumed by Hull and White (1994b), Sørensen (1994), and Chen (1996). Setting $p = 0$ gives the drift function used in those papers. The equation for x in (2.3-5) is equivalent to a functional form considered by Babbs and Webber (1994). Although set in the context of a jump model, the variable x defined in Babbs and Webber had an analogous role to that performed by x in (2.3-5). x was allowed to revert to a weighted average of μ and r to permit feedback in the economy between the economic state variable, x, and an administered short rate r. This analysis has been able to provide further justification for their assumption.

To analyse the dynamics of the system (2.3-5), it is convenient to look at the form of the matrix A when the system is expressed in the form $\dot{x} = Ax + b$. In this case (2.3-5) becomes

$$\begin{bmatrix} \dot{r} \\ \dot{x} \end{bmatrix} = \begin{bmatrix} -\alpha & \alpha \\ \beta p & -\beta \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} + \begin{bmatrix} 0 \\ \beta(1-p)\mu \end{bmatrix} \tag{2.4-1}$$

The eigenvalues of the system (2.3-5) are given by

$$\lambda = \frac{\mathrm{tr}(A) \pm \sqrt{\Delta}}{2}$$

with

$$\Delta = \mathrm{tr}(A)^2 - 4|A|$$

where A is as in (2.4-1) above. When $p > 1$ the system has two real eigenvalues, one of which will be positive so that the fixed point (r_0, x_0) is a saddle point. r may then escape to $\pm\infty$. For $1 > p \geq -(\alpha-\beta)^2/4\alpha\beta$ there are two negative eigenvalues so that the system is stable. If $p < -(\alpha-\beta)^2/4\alpha\beta$ the system has two complex eigenvalues. In this case it is always stable since $\mathrm{Re}(\lambda) = -\dfrac{\alpha+\beta}{2} < 0$ for each eigenvalue λ and the path to equilibrium will be a damped oscillatory movement. The possibilities for the dynamics may be further investigated with reference to the underlying economic parameters of the model. It is of interest whether the case of two complex eigenvalues is attainable from economic fundamentals. The term Δ determines whether the eigenvalues are complex or real. Expressing Δ for the matrix A in terms of the underlying parameters of the model, it can be seen that it can be factored in the form

$$\Delta = \frac{\left((\alpha_y - \alpha_m)u + \alpha_y(bk - uc)\right)^2}{u^2}$$

From the form of Δ above, it is the case that complex eigenvalues for the system are not obtainable based upon the underlying model formulation. If simplifying restrictions are made in the construction of (2.3-3), complex eigenvalues for the system are attainable. Notably, this may be done by assuming that y adjusts only slowly in comparison to r. This is a reasonable assumption, as speed of the money market adjustment α_m is much greater than the speed of the of the goods market adjustment α_y.

In this case, from (2.3-2) it is possible to write $dr = -\frac{1}{u}dm_d$, and the parameters of (2.3-5) reduce to

$$\begin{aligned} p &= -\frac{b}{1-c}\frac{k}{u} \\ \mu &= \frac{a - \frac{m_s}{k}(1-c)}{b + \frac{u}{k}(1-c)} \\ \alpha &= \alpha_m \\ \beta &= \alpha_y(1-c) \end{aligned} \qquad (2.4\text{-}2)$$

with the state variable x found from the transformation $x = -\frac{m_s}{u} + \frac{k}{u}y$. This form for the model does not preclude $p < -(\alpha - \beta)^2 / 4\alpha\beta$ based upon the economic formulation and hence oscillatory behaviour of the state variables is possible.

Figure 2-1 is a Monte Carlo simulation of the variable r_t in the system (2.3-5). The parameters used for the simulated path are given in Table 2-2.

TABLE 2-2: PARAMETER VALUES USED FOR FIGURE 2-1 TO FIGURE 2-6

Parameters for (2.3-5)		Underlying economic parameters		Starting values	
α	0.8	m_s	100	r_0	0.16
β	0.064	a	12.55	x_0	0.08
p	-50	b	25		
μ	0.1	c	0.6		
		u	5		
		k	4		
		α_m	0.8		
		α_y	0.16		

The underlying economic parameters may be considered to be realistic for the purposes of the example. The reversion parameters α_m and α_y control the speed with which the system reverts to equilibrium. Here, these are chosen to give realistic business cycle behaviour, taking around four years to complete a full cycle (see Figure 2-1). As posited earlier, the reversion speed for the money market is set to be faster than that for the goods market. The parameter c, being the marginal propensity to consume,. takes on a value between 0 and 1. The money supply is set to an arbitrary level, as is the parameter a representing the autonomous compoment of expenditure. The remaining parameters b, k and u reflect the determination of money demand and expenditure by the interest rate and income, as described in Section 2.3. These parameters can then be chosen such that the relative slopes of the IS and LM schedules are realistic. In an empirical analysis, Scott (1966) finds estimates for the parameters of the IS-LM model using US macroeconomic data. His findings show the IS schedule to be relatively elastic; a small drop in interest rates will lead to a large rise in expenditure (and hence income). The LM schedule is found to be relatively inelastic; a large drop in interest rates will cause only a small decrease in (speculative) money demand. To preserve equilibrium in the money market a small rise in income is required, increasing (transactions) money demand to equate money demand and supply. This empirical evidence is used in choosing the parameters b, k and u in Table 2-2 to reflect these qualitative features. Figure 2-5 and Figure 2-6 show how the choice of the paramters b, k and u translates into the IS and LM schedules.

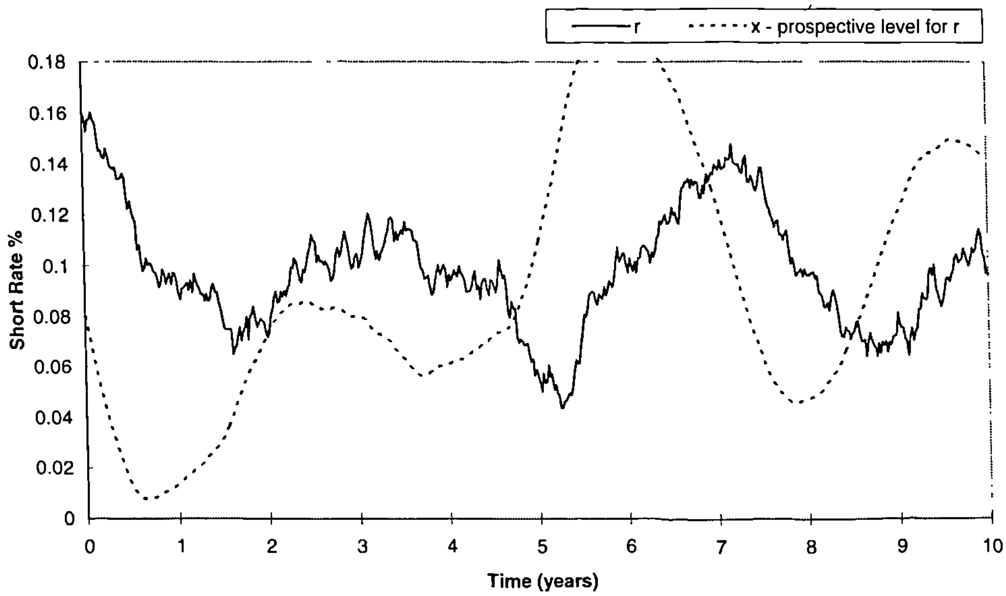


FIGURE 2-1: THE EVOLUTION OF r AND x FOR SYSTEM (2.3-5) σ_r=0.025, σ_x=0

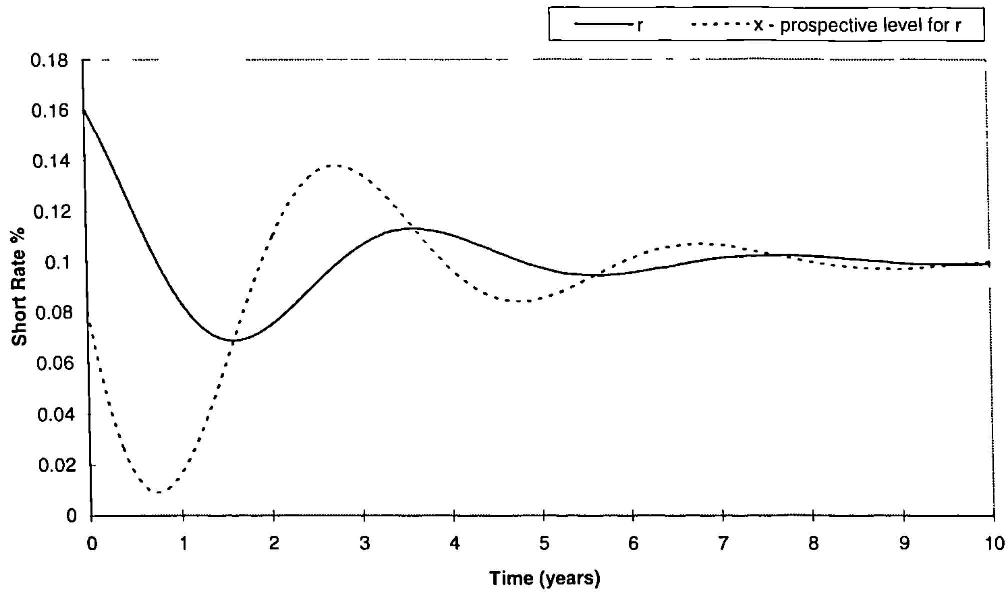


FIGURE 2-2: THE EVOLUTION OF r AND x FOR SYSTEM (2.3-5) σ_r=0, σ_x=0

The reduced form parameters in (2.3-5) are allied to the economic model via the relationships (2.4-2). r_t exhibits a strong cyclical pattern, which is evident from the fact that realistic economic parameters are chosen giving complex eigenvalues for (2.3-5). External noise is added only to the process for r. The process for x which encapsulates the economic dynamics generates the large scale fluctuations in the model.

Figure 2-2 shows the path of r_t when σ_r and σ_x are zero. This is a dampened oscillation. In the stochastic system (2.3-5) if r_t is perturbed away from mid values the dynamical behaviour of the mean causes r_t to overshoot on its return, causing oscillations. Parameter values have been fixed so that oscillation periods are several years long.[3] The qualitative dynamics of the process can be thought to be representative of business cycle type behaviour.

The oscillatory dynamics of the system for the parameter values in Table 2-2 are clearly evident if the solution path trajectory is shown in phase space. Figure 2-3 and Figure 2-4 show the phase portrait for the paths shown in Figure 2-1 and Figure 2-2 above. For the case where noise is added to the model, the path oscillates around the fix point at $r=x=\mu$. Without noise the system converges to the fix point at μ.

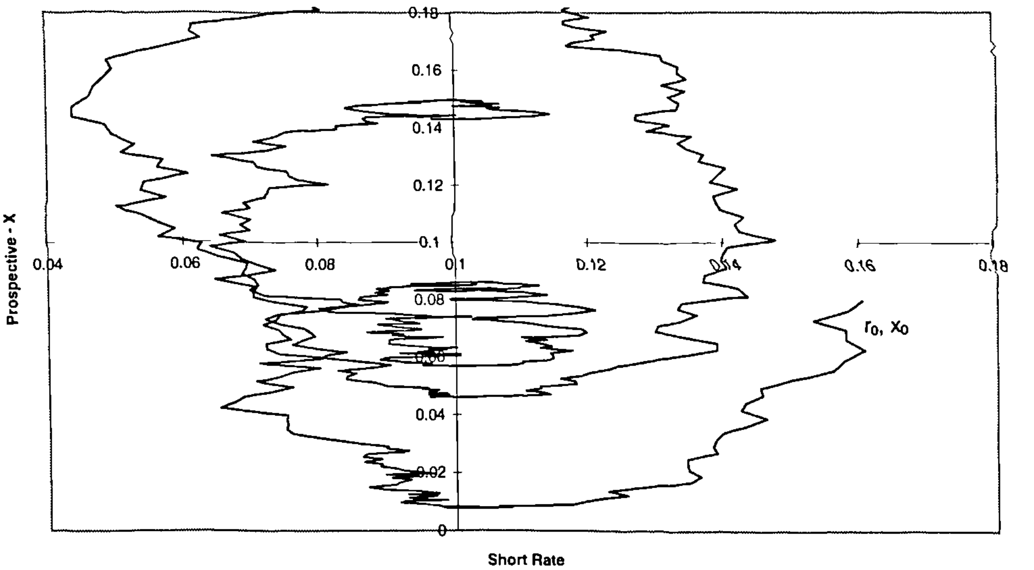


FIGURE 2-3: THE SOLUTION PATH IN r AND x SPACE σ_r=0.025, σ_x=0

[3] Beaglehole and Tenney (1991) also display an example of a model with oscillatory behaviour.

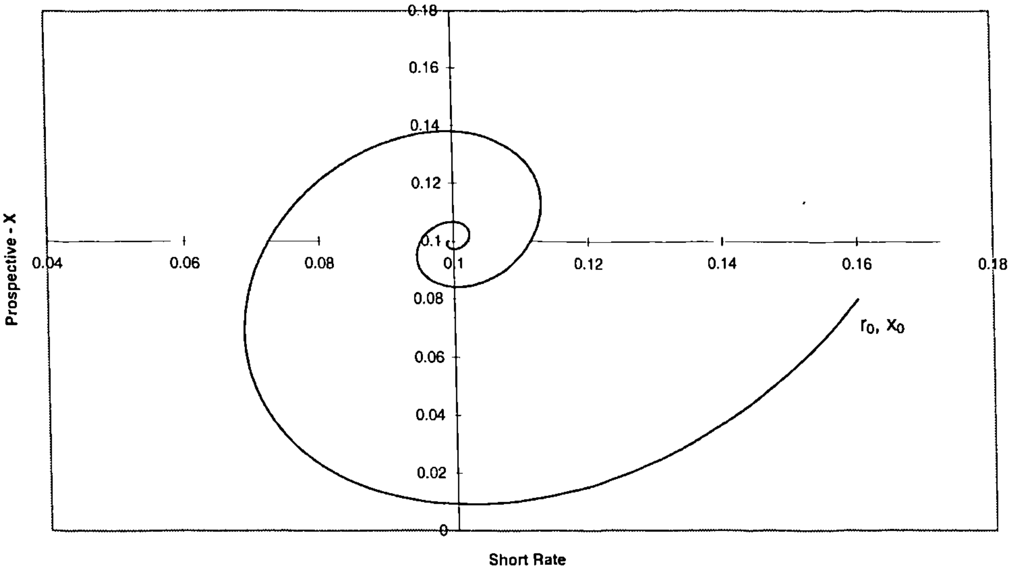


FIGURE 2-4: THE SOLUTION PATH IN r AND x SPACE σ_r=0, σ_x=0

Given the link between the reduced form model for the state variables r and x and the underlying economic model variables r and y, it is useful to view the system in the economic framework. Figure 2-5 and Figure 2-6 show the solution path in the IS-LM framework. The IS schedule represents the line, along which, the goods market is in equilibrium. For all points on the line $dy/dt = 0$. The LM schedule represents the line, along which, the money market is in equilibrium. For all points on the LM schedule, $dr/dt = 0$.

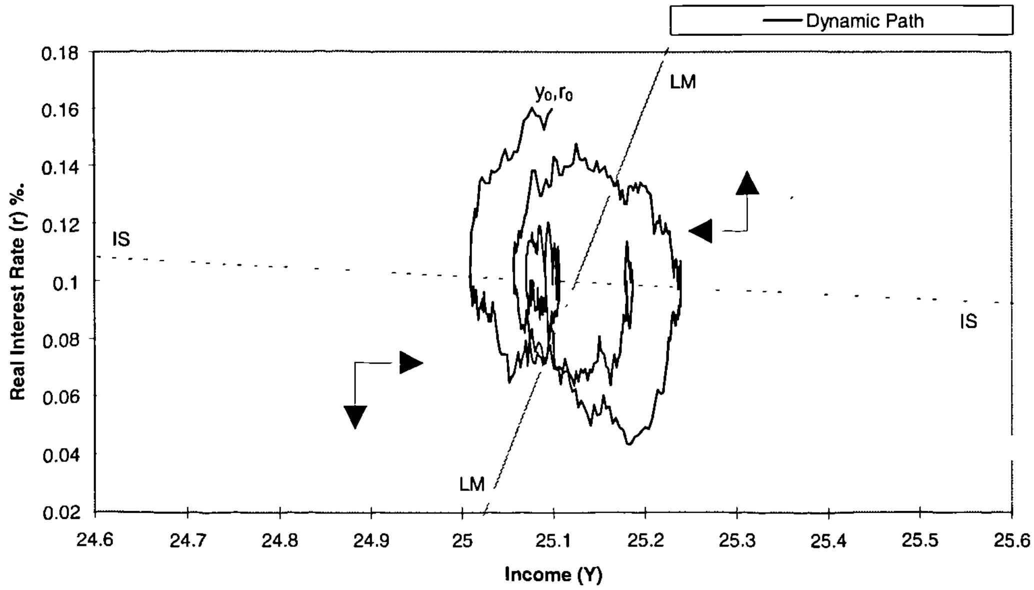


FIGURE 2-5: THE DYNAMICS OF THE SOLUTION PATH IN THE IS-LM FRAME σ_r=0.025, σ_x=0

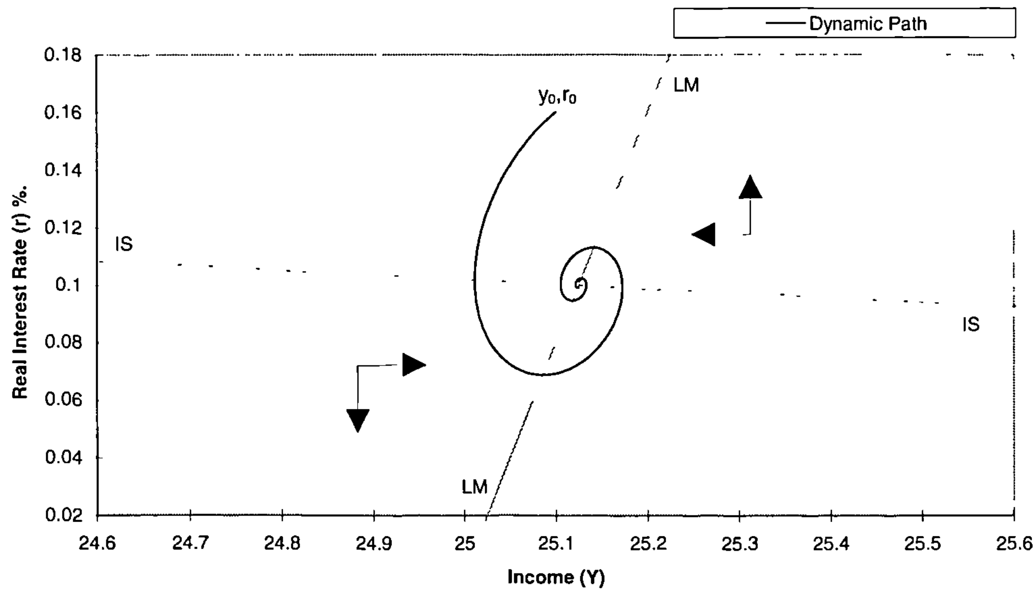


FIGURE 2-6: THE DYNAMICS OF THE SOLUTION PATH IN THE IS-LM FRAME σ_r=0, σ_x=0

Monetary authorities may influence economic variables by either fiscal or monetary means[4]. In the IS-LM framework fiscal policies determine the value of the parameter a which is assumed to incorporate taxes and government expenditure,[5] and monetary

[4] With suitable caveats. a closed economy is assumed, for instance.

[5] There are a number of alternative formulations of tax effects. Here a simple flat tax is assumed.

policies determine the value of the parameter m_s, so a and m_s may each be time dependent. Since μ depends upon a and m_s, a direct economic justification for allowing μ to be time varying can be supplied. Later, after the full model is introduced in chapter 3, alternative methods of interest rate control shall be suggested that may be available to the monetary authorities.

2.5 Conclusions

It has been seen how an economically motivated model, capable of producing complex dynamics deterministically may describe the short rate process. Furthermore, the model can be thought of as generalising several popular models in the interest rate literature. The capability of the model to describe business cycle behaviour without recourse to complex volatility functions is not without motive. The economic model, describing the interaction of the product and money markets in the economy, exhibits this behaviour endogenously for realistic parameter values. It is also the case that economic meaning can be ascribed to the reduced form parameters of the model (2.3-5) with reference to the underlying economic framework. In chapter 3 it will be seen how this model may be extended to allow more complex types of behaviour. Chapters four and five describe estimation procedures and results for the model.

3. A THREE FACTOR MODEL

3.1 Introduction

This chapter investigates a particular three factor model which, for realistic parameter values, exhibits chaotic behaviour. The three factor model arises naturally as an extension to the economically derived two factor model, shown in chapter two. The implications of its chaotic behaviour are analysed. The emphasis is concentrated on the ability of the three factor model to forge large scale deterministic dynamics. External noise serves only a minimal role in the overall dynamics. It is shown that the three factor model can replicate features of empirical interest rate processes such as business cycles and certain properties of term structures.

The chapter proceeds as follows. Section 3.2 describes an economic argument for the third factor, p, to vary. Its dynamics are motivated on behavioural grounds. A resultant three factor model is found. This then proves to be a further generalisation of the two factor model (2.3-5). The model is consistent with the description of dynamic mean models in (2.2-1). Section 3.3 discusses the implications of the form of the three factor model. It is shown that it can exhibit chaotic behaviour for realistic parameter values. Under certain transformations the model is equivalent to the Lorenz equations, the dynamics of which are well documented. One desirable property is shown to be that the system is bounded by the region of the attractor. The dynamics in the absence of noise are discussed, motivating a very simple form for the volatility functions. Section 3.4 discusses a variety of estimation and pricing issues for the model. It is shown how term structures may be obtained with a wide variety of behaviour. Bond pricing is discussed and sensitivity analysis over parameter values is presented. Discussion is made regarding the reconstruction of the attractor via principal component analysis and an empirical investigation is given, showing the existence of an attractor qualitatively similar to Lorenz. Due to the chaotic nature of the system, the stochastic version is compared to a deterministic system using an ensemble of starting states. Resultant term structures are found to be comparable. Section 3.4.2 discusses how the underlying economics of the model may imply methods for policy control of the system. Section 3.5 concludes.

3.2 Extending the two factor model

In the two factor model (2.3-5), p represents the influence of r upon the prospective level of interest rates, x. It measures the degree of feedback between r and the rest of the economy. In the derivation of (2.3-5) it was assumed that parameter values were constant, so that $p = -\frac{b}{1-c}\frac{k}{u}\frac{\alpha_m - \alpha_y(1-c)}{\alpha_m + \alpha_y bk/u}$ is constant. However, in practice parameter values, and hence the value of p, are determined by the aggregate behaviour of individuals operating within the economy. As economic activity takes place it is likely that the realised value of p will not be constant. Since α_y is small compared to α_m it is the case that $p \approx -\frac{b}{1-c}\frac{k}{u} < 0$. Macroeconomic interpretations of the parameters in this equation are discussed in chapter two. Here the interpretation is that p is related to the availability of transactions credit within the economy, via the parameter k. From the discussion in chapter two, it is was assumed that k would take on only positive values. However, economic arguments for extending the range of values that k may take on is present in the literature (for example Dornbusch and Fischer (1994), or Black and Dowd (1994)). If transaction credit is not used then k is positive. If transaction credit is available and used extensively then it is supposed that k may be negative, and hence p may be positive.

Under the supposition that k and hence p may be time varying it is necessary to model the dynamics of p. Taking equations (2.3-5) as given define

$$dp = \gamma(\bar{p} - p)dt + \sigma_p dz_p \qquad \gamma > 0, \tag{3.2-1}$$

where $\bar{p}$ is the equilibrium value of p and γ is the reversion rate of p towards $\bar{p}$. This is consistent with the general framework presented in chapter two. A detailed justification for the assumption of dynamics of the form (3.2-1) is not given, but one may suppose that as individuals become aware that the realised value of p is not in some utility maximising equilibrium, they modify their behaviour so that p is brought closer to an optimal value. An implication is that p would not revert quickly towards $\bar{p}$. Rather, it would move relatively slowly as the value of p was revealed.

It is supposed that $\bar{p}$ is a function of r and x. Expand $\bar{p}(r,x)$ in a Taylor's series expansion about the fixed point at $r_0 = x_0 = \mu$ to obtain

$$\bar{p}(r,x) = \bar{p}(\mu,\mu) + \bar{p}_r(r-\mu) + \bar{p}_x(x-\mu) + \frac{1}{2}\bar{p}_{rr}(r-\mu)^2 + \frac{1}{2}\bar{p}_{xx}(x-\mu)^2$$
$$+\bar{p}_{xr}(x-\mu)(r-\mu) + \ldots \tag{3.2-2}$$

where subscripts denote partial differentiation. It is desirable to ensure that with suitable choices of parameters r does not grow arbitrarily small or arbitrarily large. It has been seen that when $p > 1$ the system (2.3-5) goes to $\pm\infty$ (unless it lies initially on the separatrix taking it to the fixed point). Here, p represents the use of credit in the economy and an economic argument can be used to determine its behaviour. It can be expected that when both r and x, and hence r and y, are large relative to μ, credit is both costly and not needed, and p is small or negative. Conversely, when both income and interest rates are low relative to μ, despite the small cost of credit, individuals are either unwilling or unable to use credit so p is again small or negative. If both income and rates are at middle levels then credit is used. When income is high and rates are low then credit is used extensively and p may be large. However, if income is low but rates are high then credit is still used extensively. In the context of the system (2.3-5) the interpretation is for the use of credit in anticipation of interest rates decreasing.

This provides the appropriate minimal behaviour, and may be expressed by setting

$$\bar{p}(r,x) = \delta - \phi(x-\mu)(r-\mu), \qquad \phi > 0, \tag{3.2-3}$$

where $\delta = \bar{p}(\mu,\mu)$, $-\phi = \bar{p}_{xr}$ evaluated at (μ,μ), and the other derivatives have been set to zero. The form of (3.2-3) ensures that if both x and r are large or if both x and r are small, relative to μ, then p becomes small. If this happens then x begins to revert more strongly to μ than to r. Since r reverts to x, r will also begin to move back towards μ. Note that if only one of x and r is large, with the other small relative to μ, r will tend to increase if it is small and decrease if it is large. The behaviour of p does not need to be specified in this situation.

In effect, the form of (3.2-3) means that the Taylor series expansion (3.2-2) is truncated in a second order approximation.. However, the equations for r and x were obtained from the affine IS-LM system. A fully comparable system might require the development of a second order version of IS-LM. Such a system is not considered here, although one could be embedded within the framework of (2.2-3). It is the case that (2.3-5) can be regarded as a special case of a higher order approximation.[1] Note also that a purely first order approximation for $\bar{p}$ would give a three factor affine system. The dynamics of such systems (in the absence of noise) are well understood. The behaviour of r could be considered to be qualitatively similar to its behaviour in the two factor system (2.3-5).

In the analysis of the model performed below it shall be seen that p, and hence k, cycles from positive to negative values over a period of a number of years.[2] This may be interpreted as a series of credit booms and credit squeezes.

The complete three factor system is thus

[1] It is anticipated that the dynamics of a higher order approximation would not be less complex than those of the affine system presented here.
[2] Naturally, the cycle period depends on the choice of parameters.

$$
\begin{aligned}
dr &= \alpha(x-r)dt + \sigma_r dz_r \\
dx &= \beta\big(pr + (1-p)\mu - x\big)dt + \sigma_x dz_x \\
dp &= \gamma\big(\delta - \phi(x-\mu)(r-\mu) - p\big)dt + \sigma_p dz_p
\end{aligned}
\tag{3.2-4}
$$

with α large and β and γ relatively small, and the values of δ and ϕ yet to be considered. (3.2-4) is consistent with the definition (2.2-1) of a dynamic mean term structure model.

3.3 Assessing the dynamics of the model; the relation to Lorenz

The nature of the system (3.2-4) can be clarified by introducing a change of variables. For the moment, to lay bare the underlying dynamics, it is supposed that the volatility terms are identically zero so that the system is deterministic. Set

$$
\begin{aligned}
X &= \frac{r-\mu}{s} \\
Y &= \frac{x-\mu}{s} \\
Z &= \delta - p
\end{aligned}
\tag{3.3-1}
$$

where $s = \sqrt{\dfrac{1}{\gamma\phi}}$

With the assumption of zero volatilities this transforms (3.2-4) into the following system:

$$
\begin{aligned}
dX &= \alpha(Y-X)dt \\
dY &= \beta(\delta X - ZX - Y)dt \\
dZ &= (XY - \gamma Z)dt
\end{aligned}
\tag{3.3-2}
$$

With $\beta = 1$ this is the Lorenz system[3]. The Lorenz system has been widely studied, for instance see Sparrow (1982) and the references therein. For suitable values of the parameters α, δ and γ it exhibits chaotic behaviour: starting the evolution of the system from two points initially close together will produce trajectories that locally diverge exponentially while globally remaining within a bounded region. For instance, Figure

[3] β can be set to one by a suitable rescaling.

3-1[4] shows the evolution of r under two different starting conditions Path (A) has initial values (r,x,p) = (0.09, 0.085, 10) and path (B) has initial values (r,x,p) = (0.09, 0.085, 5). The paths diverge rapidly.[5] For general introductions to chaotic systems see Hilborn (1994) or Marek and Schreiber (1991), amongst others.

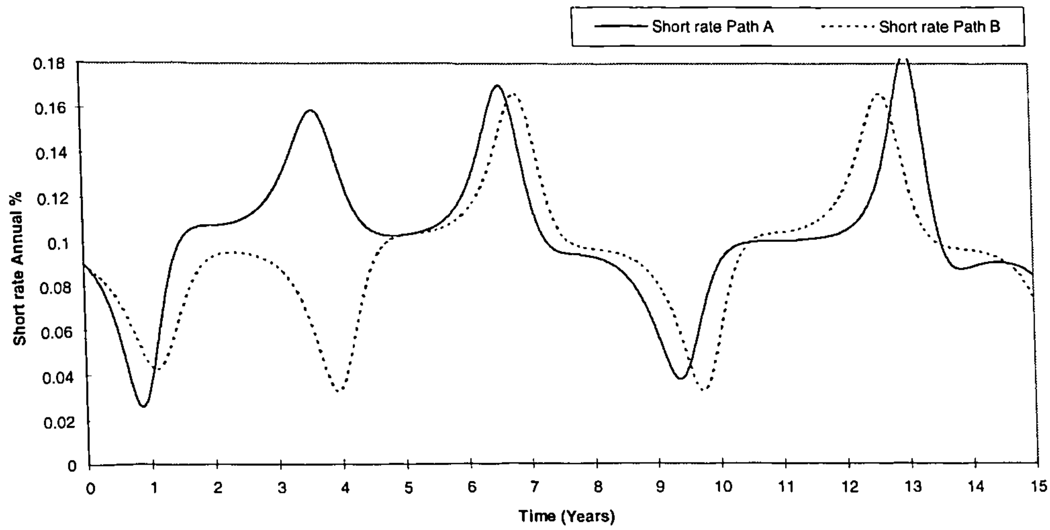


FIGURE 3-1: THE EVOLUTION OF r

In recent years there has been considerable interest in searching for non-linear and chaotic features in financial and economic time series. For instance stock market data has been examined by Abhyankar, Copeland and Wong (1997), Scheinkman and LeBaron (1992) and Hsieh (1991), and many others. For a recent review article see LeBaron (1994). Many authors attempt to uncover non-linear structure by applying one of a standard set of tests, such as the BDS test (Brock at al. (1986)) or the Grassberger-Procaccia test (1983). Barnett et al. (1995) provide a comparison of a number of tests applied to US money supply data. Fewer papers have investigated interest rate data for chaotic features. Authors who have studied T-bill data include Brock (1988), who found a correlation dimension of two, and Larrain (1991) who fitted a particular non-

[4] Numerical integration was performed using four step Runge-Kutta.
[5] The other parameter values were $\mu = 0.1$, $\alpha = 5$, $\beta = 0.5$, $\gamma = 5/12$, $\delta = 23$, $\phi = 22\times10^3$.

linear econometric model to the data. McNevin and Neftçi (1992) investigate various time series including T-bill and bond returns.

There are also comparatively few attempts to devise non-linear and chaotic models. Examples are Goodwin (1990), de Grauwe, Dewachter and Embrechts (1993), and Medio and Gallo (1992), amongst others. Goodwin explores a number of chaotic systems in economics, chiefly focusing on situations describable by the Rössler system. Chaos and business cycles have attracted some attention, for instance Brock and Sayers (1988).

A number of interest rate models incorporate some non-linearity. These include square Gaussian models (Jamshidian (1993)) and the Black-Karasinski (1991) model in which the short rate is a non-linear function of Gaussian state variables. Other models such as Longstaff (1992) and Platten (1994) incorporate non-linear terms into the drift of the short rate. Aït-Sahalia (1996) conducts an empirical investigation of Eurodollar deposit rates and concludes that the drift of the short rate is strongly non-linear. Stanton (1997), in a similar investigation using Treasury bill data reaches the same conclusion. As far as I am aware the current research presents the first example of a naturally derived non-linear model of interest rates exhibiting chaotic behaviour.

Another characteristic of the Lorenz system is also apparent in Figure 3-1. r has two ranges about which it fluctuates. Either it is oscillating with high values, greater than μ, or it is oscillating with low values, less than μ. This can be seen more clearly in Figure 3-2. This is a graph of the values of r and p as path (A) evolves. It is clear that the path loops around one of two lobes, switching from time to time from one lobe to the other. The parameters are such that switches from one regime ('high rates') to the other ('low rates') take approximately eight years, on average.

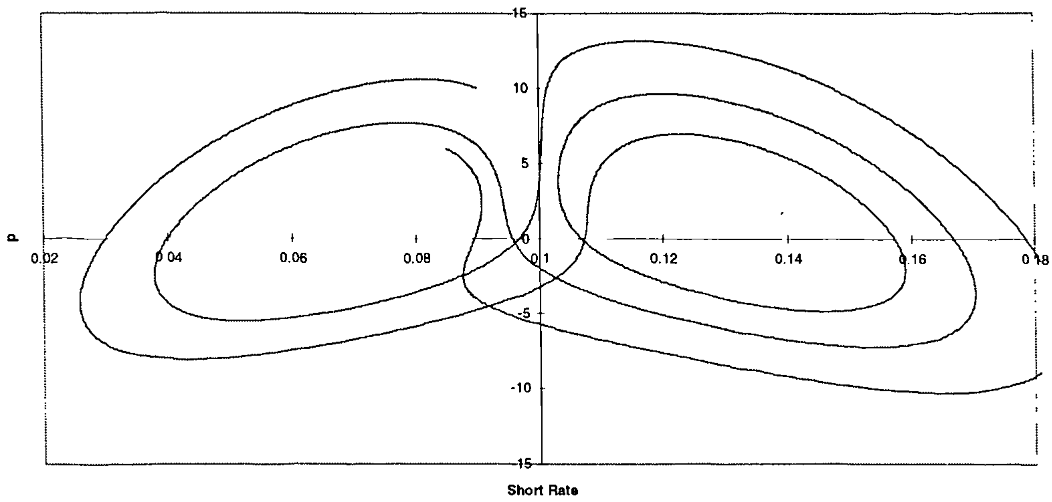


FIGURE 3-2: THE ATTRACTOR IN (r,p) SPACE

Figure 3-3 shows UK interest rates between 1954 and 1994. Although it is unlikely that a single stationary economic model would explain interest rate behaviour over this period, nevertheless there is some evidence from the figure that interest rates fall into a 'business cycle' pattern, moving from higher to lower levels and back on average every five years, or so. Figure 3-3 may be compared with Figure 3-4, obtained using illustrative parameter values. In both Figure 3-3 and Figure 3-4 interest rates are initially fluctuating at low levels. There is a sudden transition to a high rate regime, that temporarily dips back to low rates part way through. Figure 3-4 was not found through an estimation procedure, but it illustrates that the range of possible behaviours of (3.2-4) does not exclude histories qualitatively similar to the realised short rate time series.

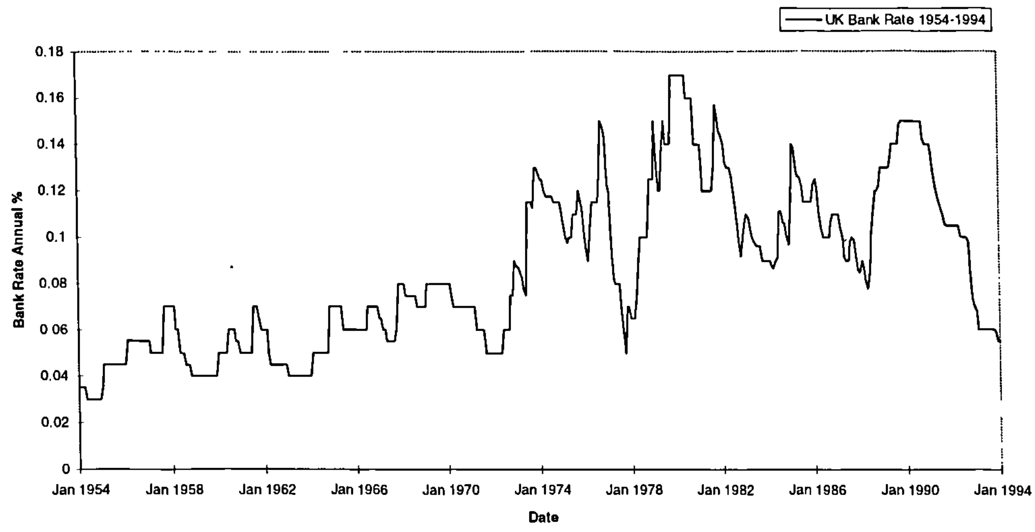


FIGURE 3-3: UK INTEREST RATE 1954-1994

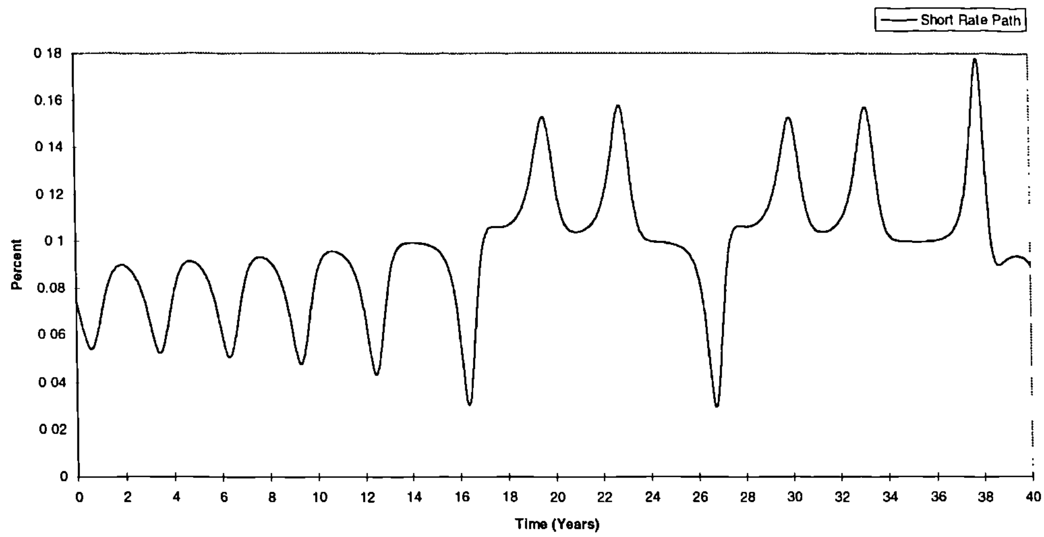


FIGURE 3-4: AN ILLUSTRATIVE EVOLUTION OF r

The behaviour of orbits in the Lorenz system (3.3-2) is highly complex, and is sensitive to the values of its parameters and to the initial state. When the parameter δ is less than 1 the system has a single fixed point at

$$X = Y = Z = 0$$

All paths converge to this fixed point, which corresponds to

$$r = x = \mu, \ p = \delta$$

When $\delta > 1$ two more fixed points exist, at

$$X = Y = \pm\sqrt{\gamma(\delta - 1)}, \quad Z = \delta - 1$$

These correspond to

$$r = x = \mu \pm \sqrt{\frac{\delta - 1}{\phi}}, \quad p = 1 \tag{3.3-3}$$

The fixed point at the origin is now a repelling fixed point. The additional fixed points may be attracting or repelling. It can be shown that if

$$\delta > \delta_H(\beta) = \frac{\alpha}{\beta}\frac{\alpha + \gamma + 3\beta}{\alpha - \gamma - \beta}$$

then each of the additional fixed points is unstable. For the illustrative example $\delta_H \approx 16.94$, so that $\delta = 23$ gives unstable fixed points.

It can also be shown that there exists an ellipsoid, containing the three fixed points, which all orbits must eventually enter and having entered cannot leave[6]. This means that, firstly, no orbits can escape to $\pm\infty$, so that r is bounded, and secondly, that when $\delta > \delta_H$ there must exist some region around which the orbits tend to be attracted. This region is the Lorenz attractor.[7] Although the structure of individual orbits about the attractor is highly complex, the geometry of the attractor itself may be significantly easier to describe, in overall terms. It shall be seen in the next chapter how the structure of the Lorenz attractor can assist attempts to estimate the system (3.2-4).

The deterministic version of (3.2-4), displays a number of characteristics that at first sight seem suited to an interest rate model. For reasonable parameter values it displays chaotic orbits around an attractor that resemble business cycle behaviour. A suitable choice of parameters can tailor the system to give cycles of various lengths.

[6] See Sparrow (1982) for a proof.

[7] The concept of an attractor can be made rigorous.

The fully stochastic system (3.2-4) has three sources of risk. In the sequel the case where both σ_x and σ_p are set to zero shall usually be considered, leaving z_r as the sole source of risk. This may be done on several grounds. Firstly, having a single source of risk allows the character of the underlying deterministic chaotic system to be foremost; the system is to be regarded as essentially a deterministic system with added noise. Secondly, it shall be seen in section 3.4.1 that the presence of noise in r is consistent with uncertainty about the value of x, and that the effect of uncertainty in x is amplified by the fact that α is relatively large. Since β and γ are both expected to be small relative to α the effect of uncertainty in x and p upon the processes for x and p will be relatively slight. Thirdly, although the results of principal component analysis of empirical term structures are often interpreted as implying that there are perhaps three sources of risk contributing to interest rate movements, it shall be indicated in section 3.4.1 that without further evidence this is entirely consistent with the evolution of a three factor dynamic mean model with a single source of risk.

Models with a stochastic mean also include noise for the process for x, and it would certainly be feasible in the three factor model to have noise both on x and on p. However, in stochastic mean models the chief effect of the noise in x is to ensure that x fluctuates sufficiently far away from its mean to generate a sufficient range of variation in r. Since in (3.2-4) the underlying attractor ensures that x and r cover a sufficient range, it is not necessary to include noise in x to produce the same effect.

The volatility term σ_r will be of Vasicek-type. This is largely for reasons of simplicity, in that allowing the stochastic term to be a function of r, for instance, would introduce an extra factor of complication. Another reason is that assuming a constant volatility is consistent with the transformation of (2.2-2) into (2.3-3). A constant volatility σ_r is consistent with constant σ_m and σ_y. Finally, although an r

dependence is introduced by some authors in order to guarantee the non-negativity of r, the presence of an attractor is relied upon to keep r positive.

The full three factor model is written as

$$\begin{aligned} dr &= \alpha(x - r)dt + \sigma dz \\ dx &= \beta\big(pr + (1 - p)\mu - x\big)dt \\ dp &= \gamma\big(\delta - \phi(x - \mu)(r - \mu) - p\big)dt \end{aligned} \tag{3.3-4}$$

where σ is a constant and z is a Wiener variable. When simulated (3.3-4) generates paths similar to Figure 3-5. Here σ has been set to 0.02, with the remaining parameters as in Figure 3-1. Figure 3-6 shows the deterministic path generated from the same starting conditions, with $\sigma = 0$. The behaviour seems qualitatively similar, in global terms. This can be seen more clearly by comparing the movements of the paths around the attractor, in Figure 3-7. The effect of noise is to cause the path of r to fluctuate somewhat as it moves around the attractor.

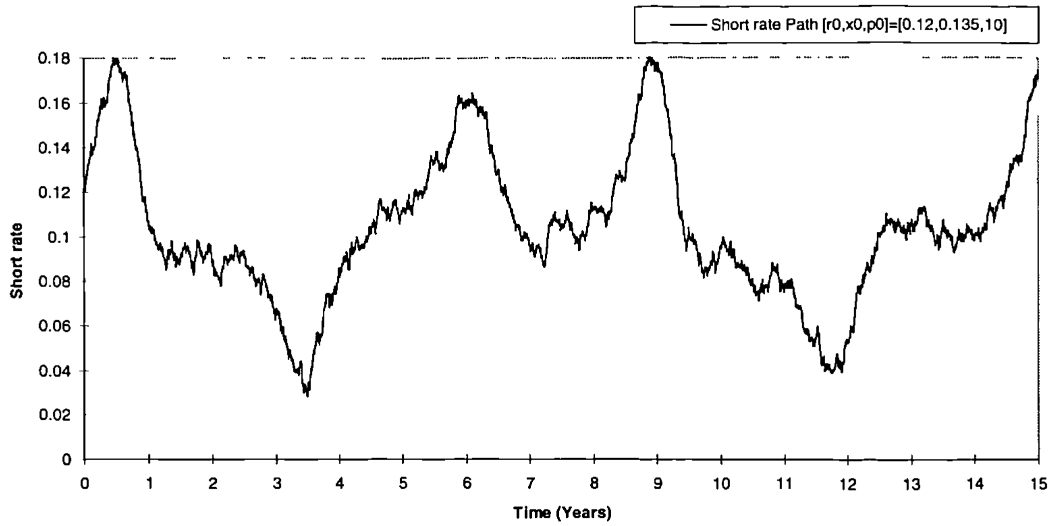


FIGURE 3-5: SHORT RATE PATH INCLUDING NOISE; σ=0.02

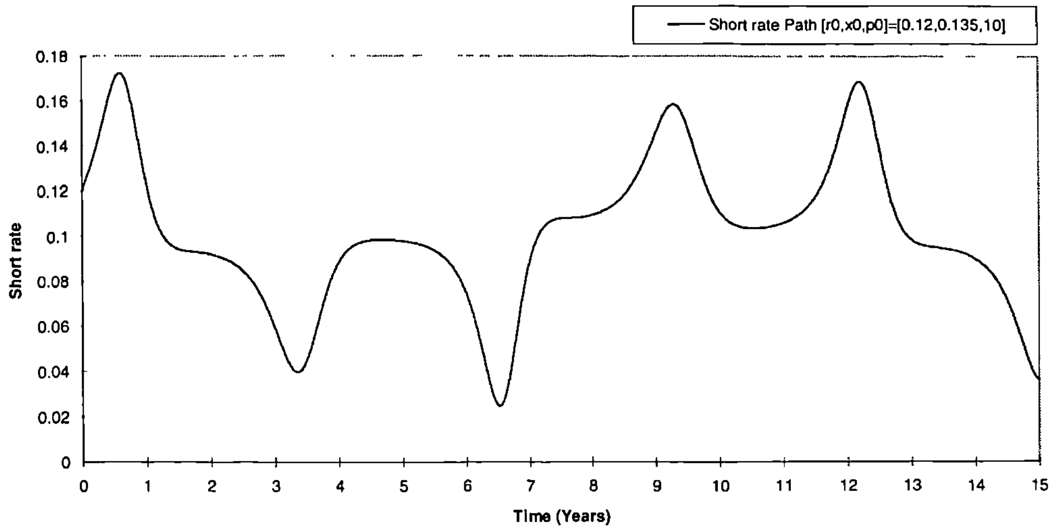


FIGURE 3-6: SHORT RATE PATH WITHOUT NOISE; σ=0

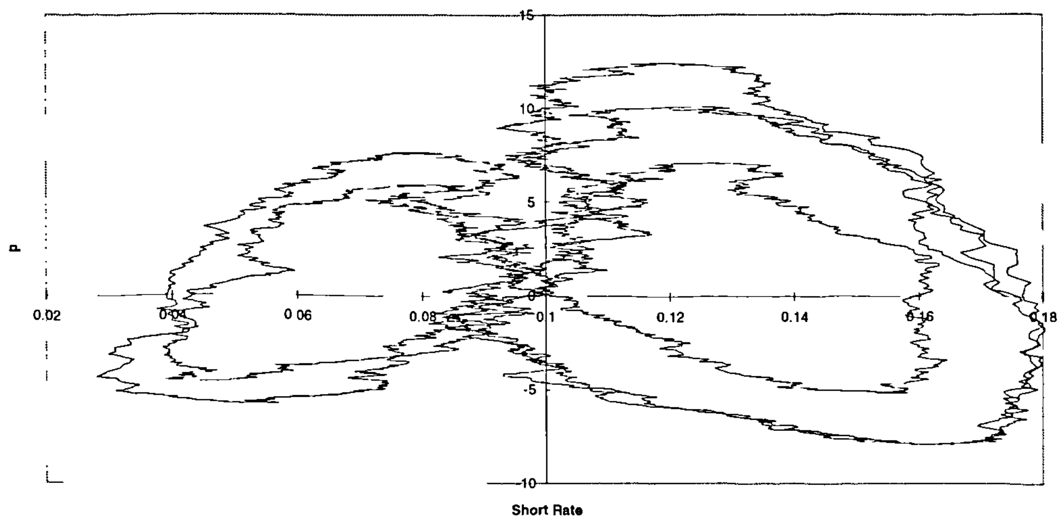


FIGURE 3-7: THE ATTRACTOR IN (r,p) SPACE WITH NOISE; σ=0.02

However great the magnitude of noise r cannot escape from the attractor. If r could escape to $+\infty$, this would amount to hyperinflation; if r went to $-\infty$, this would be another form of monetary collapse. But because the Lorenz system is bounded, all paths eventually end up at the attractor[8]. Thus the possibility of these undesirable outcomes is excluded in this model.

[8] See Sparrow (1982) for the deterministic case. The stochastic case is to be considered in the sense of Schmalfuss (1996).

The existence of an attractor is entirely analogous to the fixed point at μ in the Vasicek model, $dr = \alpha(\mu - r)dt + \sigma dz$. There, for $\sigma = 0$, every starting point converges to the fixed point. The addition of noise causes r to fluctuate about the fixed point. In this model adding noise causes r to fluctuate as it moves around the attractor.

Note the effect of the addition of a noise term in (3.3-4) is to cause r to fluctuate around a mean level x whose value, although stochastic because of its dependence on r, has relatively little noise. It is as if r reverts to a deterministic reversion level, but with feedback between the reversion level and the value of r. This is consistent with the findings of Chan, Karolyi, Longstaff and Sanders (1992) reported in the introduction.

3.4 Estimation and pricing issues

Write $B_t(T)$ for the value at time t of a pure discount bond yielding 1 at time T, and define $r_t(T)$ as the spot rate corresponding to $B_t(T)$,

$$B_t(T) = \exp[-(T,t)\cdot r_t(T)]$$

If the value of $B_t(T)$ is contingent upon the values of r, x and p then $B_t(T) = B_t(T, r, x, p)$ satisfies the usual bond pricing equation:

$$rB = \frac{1}{2}\sum_{i,j=r,x,p} \rho_{ij}\sigma_i\sigma_j \frac{\partial^2 B}{\partial i \partial j} + \sum_{i=r,x,p} (\mu_i - \lambda_i\sigma_i)\frac{\partial B}{\partial i} + \frac{\partial B}{\partial t} \qquad (3.4\text{-}1)$$

where μ_i and σ_i are the drifts and volatilities of r, x and p, ρ_{ij} is the correlation between z_i and z_j, the λ_i are the prices of risk for r, x and p, and the boundary condition is $B_T(T) = 1$. In general this equation may only be solved numerically, but it is feasible to obtain term structures using Monte Carlo methods.

Figure 3-8 shows a term structure of interest rates computed by a Monte Carlo method under an assumption of risk neutrality[9]. The initial state was $(r_0, x_0, p_0) = (0.09,$

[9] The assumption of risk neutrality is not crucial.

0.085, 10), with the usual parameter values. For simplicity it was assumed that $\sigma_x = \sigma_p = 0$, and $\sigma_r = 0.02$. This is equivalent to solving the partial differential equation

$$\begin{aligned} rB = {} & \frac{1}{2}\sigma_r^2\frac{\partial^2 B}{\partial r^2} + (\alpha(x-r) - \lambda_r\sigma_r)\frac{\partial B}{\partial r} + \beta(pr + (1-p)\mu - x)\frac{\partial B}{\partial x} \\ & + \gamma(\delta - \phi(x-\mu)(r-\mu) - p)\frac{\partial B}{\partial p} + \frac{\partial B}{\partial t} \end{aligned} \quad (3.4\text{-}2)$$

with $\lambda_r = 0$. The only second order term is in r.

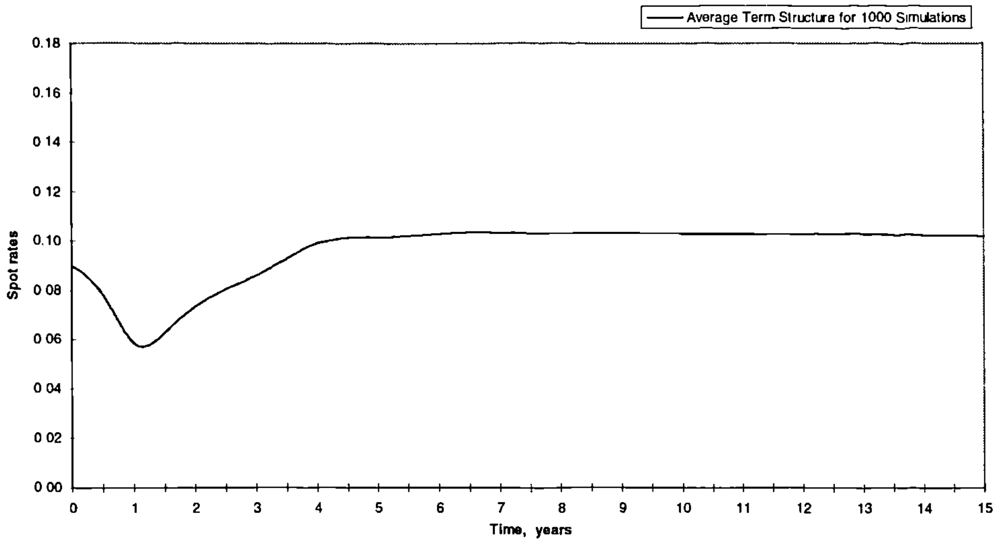


FIGURE 3-8: THE TERM STRUCTURE OF INTEREST RATES: A MONTE CARLO ESTIMATE

The behaviour of (3.4-2) is not investigated here other than through Monte Carlo analysis. Monte Carlo methods seem to be effective at solving for bond prices, despite the chaotic nature of the system, because of the time scale of the system.[10] Since cycles in the model are of three to five years duration, evolving term structures over a period of twenty or more years involves comparatively few cycles. The Lorenz system is relative stable over a small number of cycles; it was found by experimenting with different numerical algorithms that computational inaccuracies were ignorable, and slight changes in initial conditions did not often greatly alter the evolution of the short rate until at least a significant fraction of a cycle had elapsed, in the worst case.

[10] However, it is possible that solving the partial differential equation directly may produce results different to those found by Monte Carlo methods.

The main problem likely to be faced in practice in fitting the system (3.3-4) to bond prices is the difficulty in estimating parameter values and the initial values of the system. Preliminary sensitivity analysis seems to indicate that bond and bond option prices (which may also be computed by Monte Carlo methods) are relatively insensitive to changes in parameter values. Some results are presented in Table 3-1 and Table 3-2. For the chosen initial values bond option prices were least sensitive to changes in σ, λ and β, more sensitive to changes in γ and ϕ, and most sensitive to changes in α, δ and μ, with μ being significantly more important than the others. Chapters four and five describe methods that may be used for estimating the parameters of the model.

TABLE 3-1: SENSITIVITY OF BOND OPTION PRICES TO PARAMETER VALUES

Percentage change in bond option value under various percentage changes to the parameters.

Base Case Monte-Carlo Value of Call : 0.042
Standard error of estimate : *0.0003*

		Percentage change in Parameter			
Parameter	Base Value	-5%	-1%	+1%	+5%
α	5	3.96	0.76	-0.73	-3.53
δ	23	2.16	0.42	-0.51	-3.24
γ	5/12	-0.73	-0.21	0.39	0.45
μ	0.1	38.3	7.53	-7.41	-35.76
ϕ	22×10^3	1.40	0.28	-0.30	-1.45
β	0.5	-0.16	0.05	0.12	-0.31
σ	0.025	-0.11	-0.02	-0.03	-0.19
λ	0.1	-0.25	-0.03	0.05	0.24

Prices computed using a Monte Carlo method with antithetic variables and 10,000 sample paths per estimated call price.
The bond option is a two year European option on a ten year bond, with an exercise rate of 10%

TABLE 3-2: SENSITIVITY OF BOND PRICES TO PARAMETER VALUES

Percentage change in price of a 2 yr bond under various percentage changes to the parameters.
Base Case Monte Carlo Estimate of Bond Price $B_0(2)$ 0.788
Standard Error of Estimate *0.00004*

		Percentage change in Parameter			
Parameter	Base Value	-5%	-1%	+1%	+5%
α	5	0.02	0.00	0.00	-0.02
δ	23	0.00	0.00	0.00	-0.01
γ	5/12	-0.15	-0.03	0.03	0.14
μ	0.1	1.23	0.25	-0.24	-1.22
ϕ	$22\text{x}10^3$	-0.16	-0.03	0.03	0.15
β	0.5	-0.18	-0.04	0.04	0.17
σ	0.025	0.00	0.00	0.00	0.00
λ	0.1	0.00	0.00	0.00	0.00

Percentage change in price of a 5 yr bond under various percentage changes to the parameters.
Base Case Monte Carlo Estimate of Bond Price $B_0(5)$ 0.620
Standard Error of Estimate *0.0002*

		Percentage change in Parameter			
Parameter	Base Value	-5%	-1%	+1%	+5%
α	5	0.23	0.04	-0.05	-0.26
δ	23	-0.17	-0.03	0.02	0.11
γ	5/12	-0.23	-0.04	0.04	0.18
μ	0.1	2.91	0.58	-0.59	-2.94
ϕ	$22\text{x}10^3$	-0.01	0.00	0.00	0.00
β	0.5	-0.32	-0.06	0.06	0.26
σ	0.025	0.09	0.02	-0.03	-0.13
λ	0.1	-0.01	0.00	0.00	0.01

Percentage change in price of a 10 yr bond under various percentage changes to the parameters.
Base Case Monte Carlo Estimate of Bond Price $B_0(10)$ 0.392
Standard Error of Estimate *0.0004*

		Percentage change in Parameter			
Parameter	Base Value	-5%	-1%	+1%	+5%
α	5	0.51	0.09	-0.10	-0.48
δ	23	0.23	0.03	-0.06	-0.36
γ	5/12	-0.24	-0.06	0.07	0.19
μ	0.1	5.82	1.16	-1.15	-5.69
ϕ	$22\text{x}10^3$	-0.01	0.00	0.00	-0.01
β	0.5	-0.27	-0.05	0.05	0.18
σ	0.025	0.07	0.01	-0.02	-0.12
λ	0.1	-0.03	0.00	0.01	0.03

Prices computed using a Monte Carlo method with antithetic variables and 10,000 sample paths per estimated bond price.

There is a wide literature on methods of model and parameter estimation in non-linear and chaotic dynamical systems. See for instance the review article by Abarbanel et al. (1993), and for a comparison of tests see Barnett et al (1995). Important factors that contribute to the ease or difficulty of estimation are:

i) whether observations are noisy or 'clean'

ii) whether the data series is long or short, relative to the cycle time of the system.

iii) whether spatial data is available.

For interest rate data there is relatively little noise compared to some other physical systems. It is arguable that the presence of noise may actually be helpful to estimation. Furthermore, spatial data is available in the form of spot rates over various times to maturity, not just the short rate itself. The greatest impediment to estimation is the relative paucity of time series data. Although several thousand daily observations were used to produce Figure 3-3, this represents only a few cycles around the attractor of (3.3-4). Much effort in the empirical investigation of non-linear systems has gone into either detecting non-linear behaviour, or into attempting to specify a non-linear model. For the latter task in particular a huge quantity of time series data is required to ensure that any underlying attractor is sufficiently covered by sample points. Fortunately, the problem here is different. The objective is simply to estimate parameters of the system specified by (3.3-4). For this task much less data is required. To some extent it may be tackled by exploiting the known geometry of the Lorenz attractor, and the presence of spatial data. Techniques to estimate the parameters of the system are presented in chapters four and five.

The presence of spatial data in the form of term structures enables us to investigate non-linearities inherent in interest rate data. One method of using spatial data, as described in Abarbanel et al. (1993), Broomhead and King (1986) and Kostelich (1992), is to investigate the principal components contributing to the time evolution of the spatial data[11,12]. In this case, the method of singular-value analysis corresponds to finding functions $\sigma^i(\tau)$, $i = 0,\ldots,n$, such that

$$r_t(t+\tau) = \sigma^0(\tau) + \sum_{i=1}^{n} \sigma^i(\tau) z^i(t) + \varepsilon_t(\tau) \tag{3.4-3}$$

The numbers $z^i(t)$, $i = 1,\ldots,n$, measure the size of the contribution of the eigenfunctions $\sigma^i(\tau)$ to the term structure at time t. $\varepsilon_t(\tau)$ represents residual error. From the data an n can be found such that n components $\sigma^i(\tau)$ are significant, and no other significant components can be detected. n specifies the embedding dimension of the system. It is the minimum dimension of a vector space into which an underlying attractor may be embedded. The evolution of the system (3.4-3) is equivalent to the evolution of the vector $z_t = \{z^i(t)\}_{i=1,\ldots,n}$, in the embedding space. In a deterministic system the evolution of z_t will be strongly related to that of the underlying dynamical system that generated the spatial data $r_t(t+\tau)$. In a stochastic system z_t will be a random variable.

Note that the decomposition (3.4-3) is related to the evolution of spot rates in a Heath, Jarrow and Morton model. In practical implementations of Heath, Jarrow and Morton type models volatility functions $\sigma^i(\tau)$ are found from historical term structure data, or from the prices of other term structure dependent instruments, by principal component methods.[13] In such a model z_t is presumed to be a vector of independent standard Wiener processes. However, the differenced series of empirically obtained z_t

[11] An application to the Lorenz system is given in Rowlands and Sprott (1992).
[12] For an alternative approach see Kwasniok (1996).
[13] Strictly, the components in an HJM model are components of the forward rate curve. In practice, components are often obtained from spot rates or directly from bond data.

are likely to exhibit striking serial correlation and heteroskedasticity.[14] This is evidence of underlying dynamical structure, and suggests that the series z_t be analysed with the objective of uncovering such structure. Such an investigation is likely to prove far more fruitful than an analysis based on short rate data alone, since the information content of the term structure is much greater than that of a single spot rate. For instance, based on principal component analysis it is frequently reported that up to three factors are sufficient to 'explain' interest rate dynamics (as one example see Steeley (1989)). This is simply saying that interest rates are a low-dimensional system with embedding dimension equal to or less than three. It would be difficult to reach this conclusion solely from the analysis of short rate data.

An analysis of a series of z_t for sterling data has been carried out by Nunes and Webber (1996). They find intriguing evidence of structure in their sample. Here, that analysis is replicated for the purpose showing how the structure found may be thought of as comparable to an attractor in the embedding dimension.

PCA analysis is conducted on UK money market data from 28/09/88 to 03/02/95[15]. In total 1609 daily observations are available for three, six and twelve month LIBOR as well two, three, four, five, seven and ten year swap rates. In (3.4-3) above, this gives n=9. Consistent with Nunes and Webber (1996), the data is not stripped, but market rates are used directly[16]. Further, to remove jump elements from the data, outliers are removed[17]. The results for the analysis are shown in Table 3-3 below.

[14] Private communication, Les Clewlow, University of Warwick. See also Nunes and Webber (1996).
[15] Thanks go to Les Clewlow for provison of the data for this analysis.
[16] Stripping is not performed to avoid biasing the results through the particular stripping method. It should, as such, be recognised, that the components and dynamics explain the quoted market rates, rather than spot rates.
[17] Nunes and Webber (1996) remove 16 outliers. I have chosen to remove only the three most significant outliers.

TABLE 3-3: PCA RESULTS FOR UK MONEY MARKET DATA[18]

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8	σ^9
E-value	0.06259	0.01704	0.01125	0.00393	0.00250	0.00113	0.00042	0.00027	0.00023
Contribution	*0.62996*	*0.17150*	*0.11319*	*0.03957*	*0.02515*	*0.01138*	*0.00423*	*0.00269*	*0.00234*
Cumulative	*0.62996*	*0.80146*	*0.91464*	*0.95421*	*0.97936*	*0.99074*	*0.99497*	*0.99766*	*1.00000*
					Eigenvectors				
3 mth	-0.56774	0.67997	0.42661	0.14814	0.05365	0.09206	-0.00281	-0.00225	-0.00172
6 mth	-0.33812	0.13656	-0.58608	0.01617	-0.72312	0.01731	-0.00511	0.00507	-0.00504
1 yr	-0.34225	0.04998	-0.63741	0.04230	0.68716	0.00515	-0.00710	-0.00372	0.00259
2 yr	-0.34729	-0.21276	0.12441	-0.86670	0.00841	0.23944	-0.09947	-0.01116	-0.00730
3 yr	-0.30024	-0.23597	0.11234	-0.03916	-0.01399	-0.53214	0.73811	0.08090	0.07280
4 yr	-0.27946	-0.28485	0.11668	0.14230	-0.02010	-0.41000	-0.41320	-0.38005	-0.56852
5 yr	-0.26554	-0.30057	0.11065	0.18230	-0.02401	-0.20618	-0.44809	0.14623	0.72692
7 yr	-0.21993	-0.34285	0.09110	0.27176	-0.01414	0.33568	-0.00044	0.71683	-0.35384
10 yr	-0.19615	-0.35570	0.07727	0.30978	-0.02423	0.57219	0.27146	-0.56002	0.13331

Consistent with Nunes and Webber (1996), the first three factors account for the majority of the variation in the term structure. The first factor accounts for by far the largest share, of 63%. The second and third factors account for 17% and 11% respectively. In total, the first three factors account for 91.5% of the variation in the term structure. It is concluded, on the basis of Table 3-3, that three factors are significant. This judgement is not made on rigorous grounds, but is in agreement with Nunes and Webber (1996) and Steeley (1989). Here, the interest is in the structure of $z(t)$. Plotting the elements of $z^i(t)$ for i=1,2,3 can be thought of as reconstructing the attractor in the embedding space. If no attractor exists, then the elements of $z(t)$ will be evolve as a random walk. This may likely be driftless, and without any apparent structure. Figure 3-9 and Figure 3-10 below present two perspectives of the elements of $z(t)$ in three space.

[18] Values differ marginally from those presented in Nunes and Webber due to differing numbers of outliers having been removed.

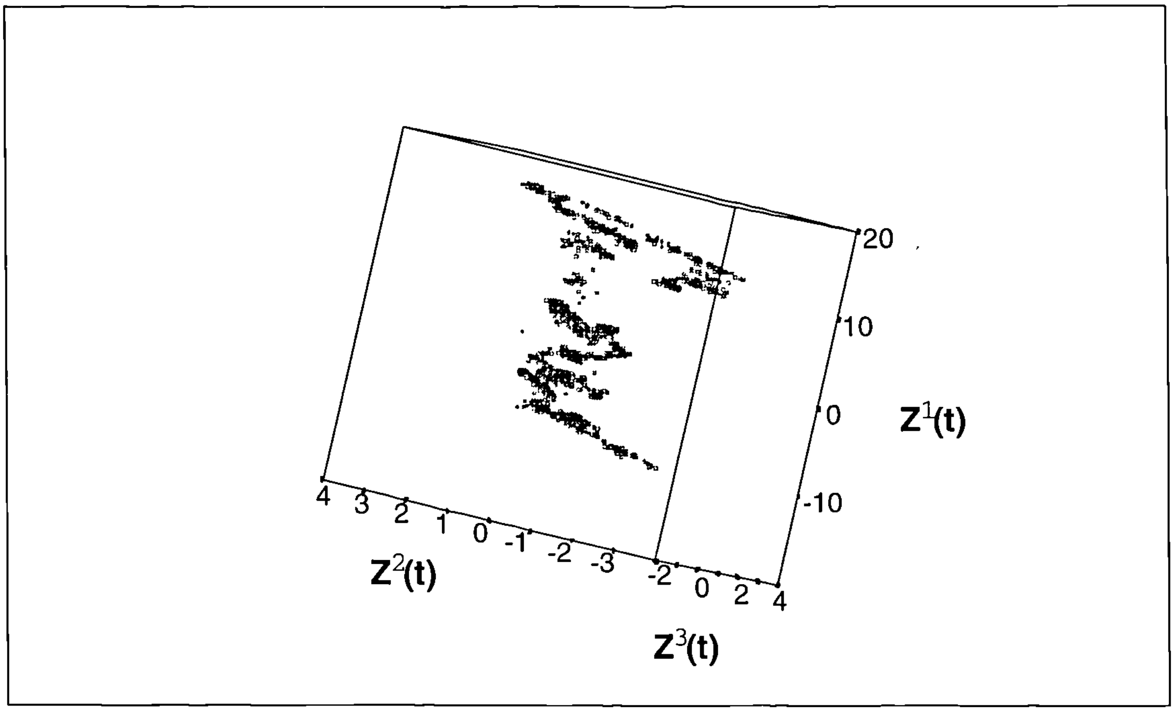


FIGURE 3-9: PERSPECTIVE OF z_t FOR n=3

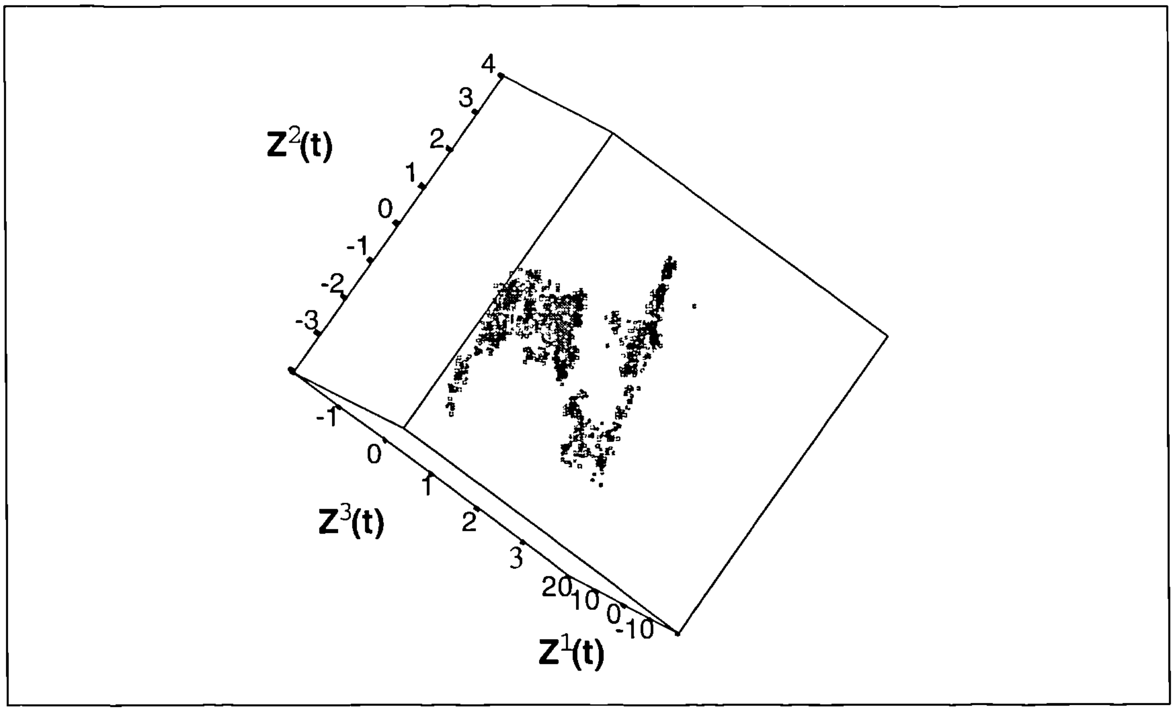


FIGURE 3-10: SECOND PERSPECTIVE OF z_t FOR n=3

From looking at Figure 3-9 and Figure 3-10 it is clear that z_t is highly structured and certainly not following a random path over the space. Indeed it is possible, to make out the presence of what could be a two lobed attractor. Qualitatively, this is the sort of dynamics the system (3.2-4) can be expected to produce. Given that the data set here covers only some six and half years, this will likely represent no more than two cycles of the attractor. It is then not to be expected that the series z_t will accurately reproduce the full extent of the structure that the process possesses. While this is not a rigorous exposition of the factors involved here, it certainly provides evidence for the belief that z_t is not a purely stochastic process but rather that there are deterministic elements involved.

Note that the existence of three components does not necessarily imply the existence of three sources of risk; without further evidence it is just a statement about the embedding dimension.

It is noted in passing that spatial data exists in options markets in the form of option prices at different strikes and different times to maturity. This applies to equity and foreign exchange markets as well as interest rate markets. Such data can be analysed as indicated, and may give insights into the dynamics of particular markets.

3.4.1 Observability, ensembles and noise

If the deterministic version of (3.2-4) were to be used as a basis for pricing financial instruments it would not be strictly necessary to introduce a market price of risk. For instance, should one wish to calculate the term structure of interest rates from the deterministic version of (3.2-4) one would have,

$$B_t(T) = \exp\left[-\int_t^T r(u)du\right] \tag{3.4-4}$$

and spot rates, $r_t(T)$, would be obtained from bond prices as usual:

$$r_t(T) = \frac{1}{T-t}\int_t^T r(u)du \tag{3.4-5}$$

Figure 3-11 shows the term structure of interest rates derived from paths (A) and (B) in the deterministic version of (3.2-4). The time evolution of the term structure is straight forward; one simply starts the integral (3.4-5) at a different point along the path. Figure 3-12 shows the evolution of term structures resulting from the path shown in Figure 3-13. Figure 3-13 shows the short rate as it shifts from a high regime to a low regime, resulting in the dramatic term structures of Figure 3-12.

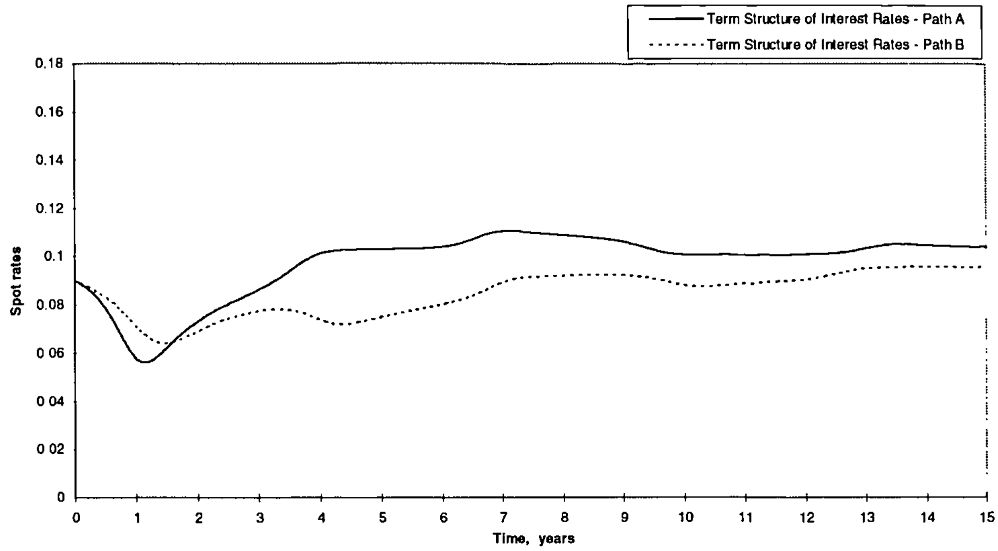


FIGURE 3-11: DETERMINISTIC TERM STRUCTURES FOR PATHS A AND B

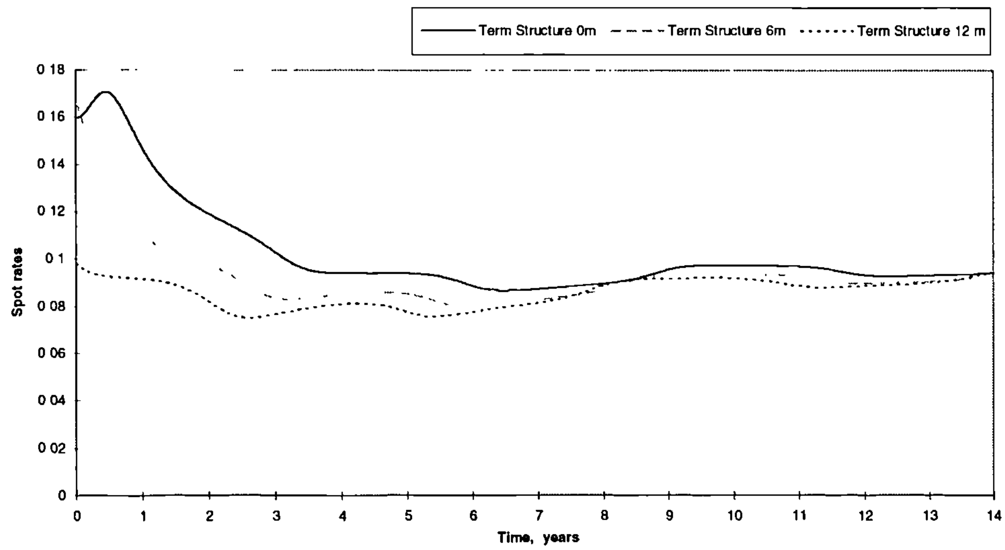


FIGURE 3-12: EVOLUTION OF THE TERM STRUCTURE

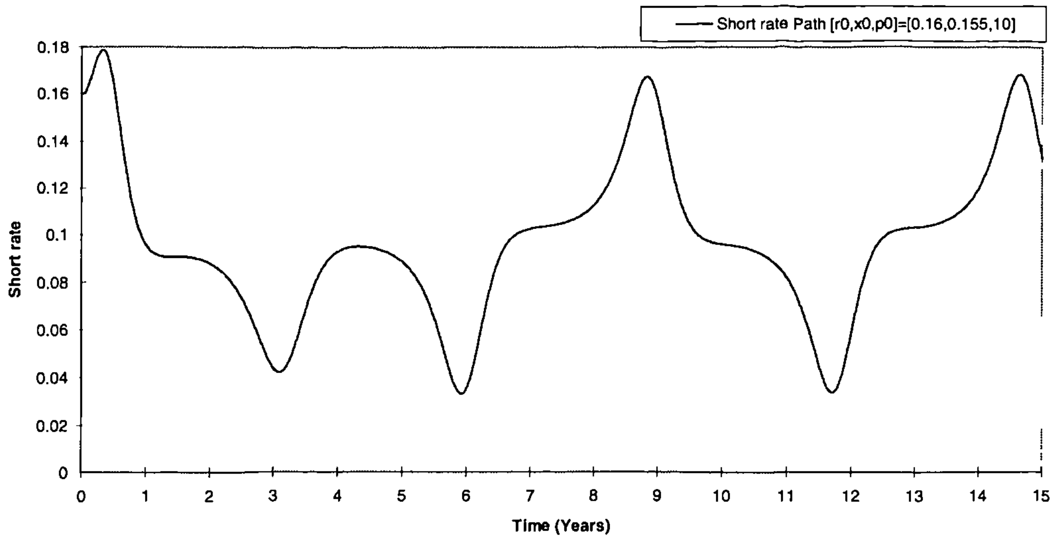


FIGURE 3-13: DETERMINISTIC SAMPLE PATH FOR FIGURE 3-12

However, (3.2-4) is a chaotic system. If it is not possible to observe with certainty the initial values of the system, the future state of the system is essentially unpredictable. In this context the current values of the variables x and p are not directly observed, and may only be inferred. One is led to consider the evolution of an ensemble of states, leading to a probability distribution over future states (as in, for instance, Smith (1995)).

For most models, being unable to exactly specify the current state is not a matter of any particular importance. With zero volatility the Vasicek system, for instance, rapidly converges to a fixed point, and with non-zero volatility any slight uncertainty about the current state is immediately obscured by noise. With a dissipative system such as (3.2-4), however, it is necessary to adopt an ensemble approach. Instead of evolving the system from a single initial state (r_0, x_0, p_0), one considers the time evolution of a neighbourhood of states.

Figure 3-14 shows the distribution of terminal values of r after evolving the system for 15 years from an initial state of (0.16, x_0, 1), where x_0 runs between 0.14 and 0.17 in steps of 0.00003. If the perceived value of x were uniformly distributed in the range [0.14,0.17], the distribution of terminal values of r might resemble the figure.

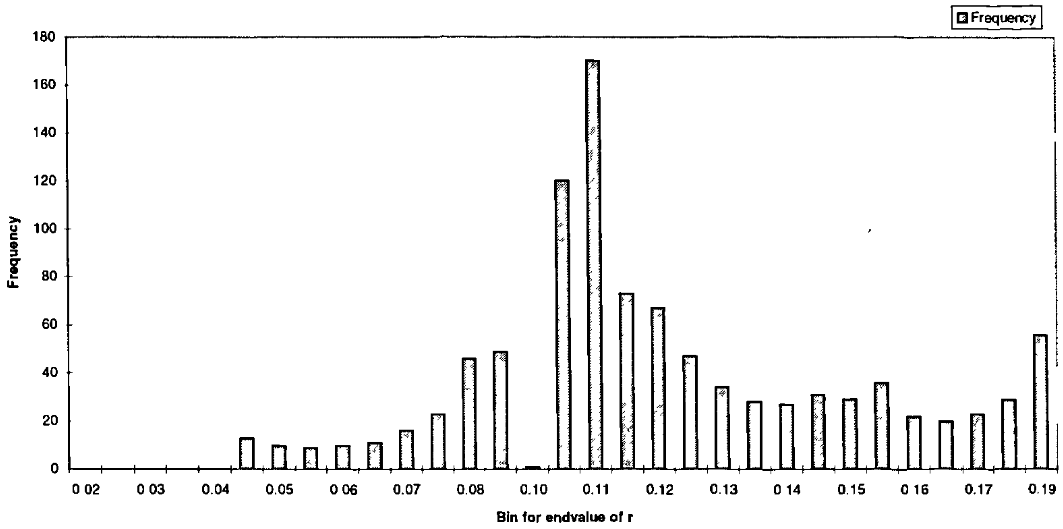


FIGURE 3-14: THE DISTRIBUTION OF TERMINAL VALUES FOR r: AN ENSEMBLE ESTIMATE

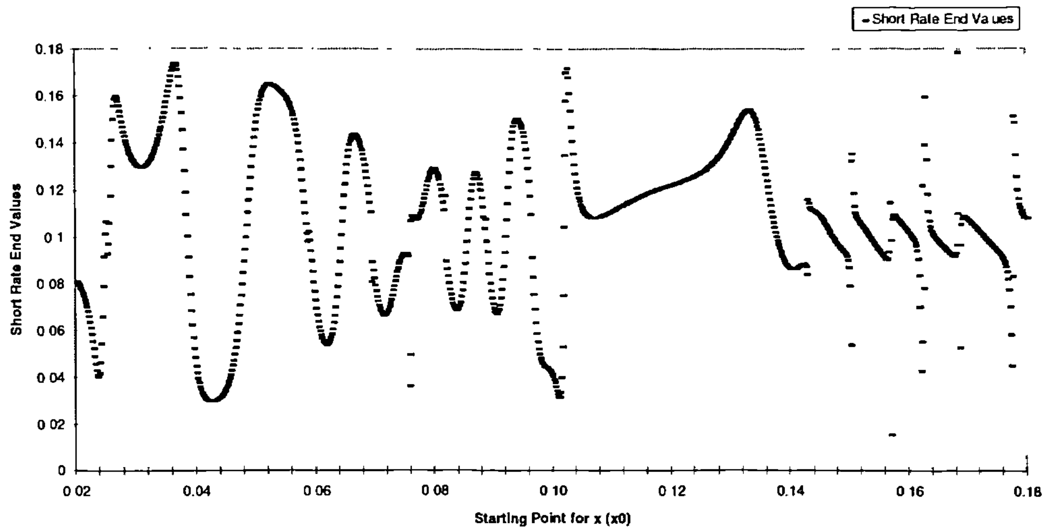


FIGURE 3-15: TERMINAL VALUES FOR r: x_0 IN THE RANGE [0.02,0.18]

An example of the dependence of r upon x can be seen in Figure 3-15[19], where terminal values of r have been plotted against initial values of x for initial values (0.16, x_0, 1), x_0 in the range [0.02, 0.18]. The variation in end values of r is dramatic.

[19] Parameter values used for this figure are as follows : α=2.5, β=0.5, γ=0.208, δ=20, ϕ=42660, μ=0.1. $\delta>\delta_H$=16.94 (condition for chaos satisfied)

Although the end value of r is a smooth function of x there are regions where very rapid switches take place.

The distribution of end values of r may also be estimated when the initial state is fixed but r has a stochastic dependence. Figure 3-16 shows the distribution of 1000 simulations of the evolution of r under (3.3-4) using the usual parameter values and σ=0.02. Although less bunched than the ensemble distribution of Figure 3-14, the distributions are nevertheless not incomparable.

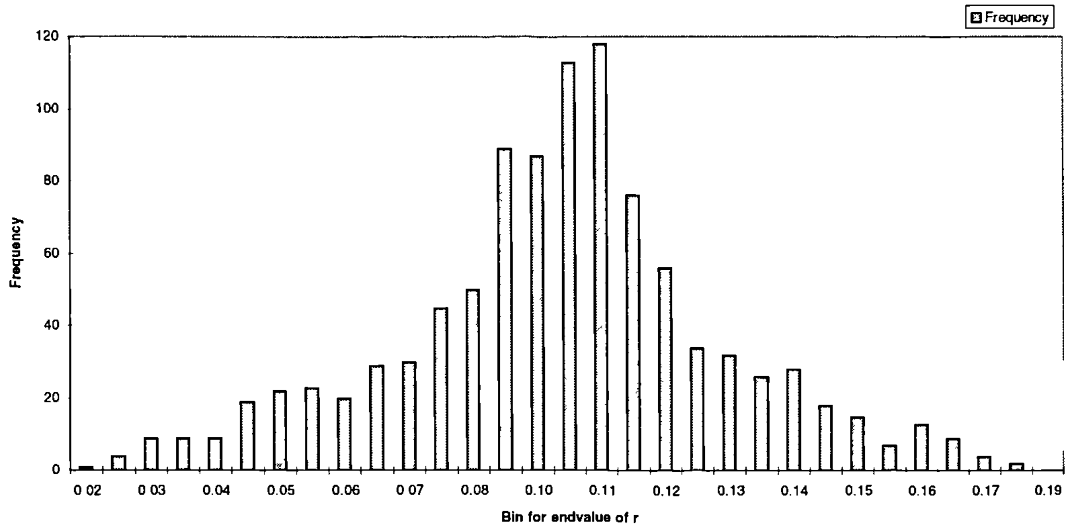


FIGURE 3-16: THE DISTRIBUTION OF TERMINAL VALUES FOR r: A STOCHASTIC ESTIMATE

Figure 3-17 shows a term structure computed using an ensemble approach[20]. This can be compared to Figure 3-8. The term structures are remarkably similar. Both of these figures are similar to Figure 3-11 which shows the deterministic term structure computed from an initial state $(r_0, x_0, p_0) = (0.09, 0.085, 10)$.

[20] Figure 3-17 was produced by computing deterministic bond prices for each x0 in the range [0.08, 0.09], in steps of 0.00001, with r_0 and p_0 kept fixed at 0.09 and 10 respectively.

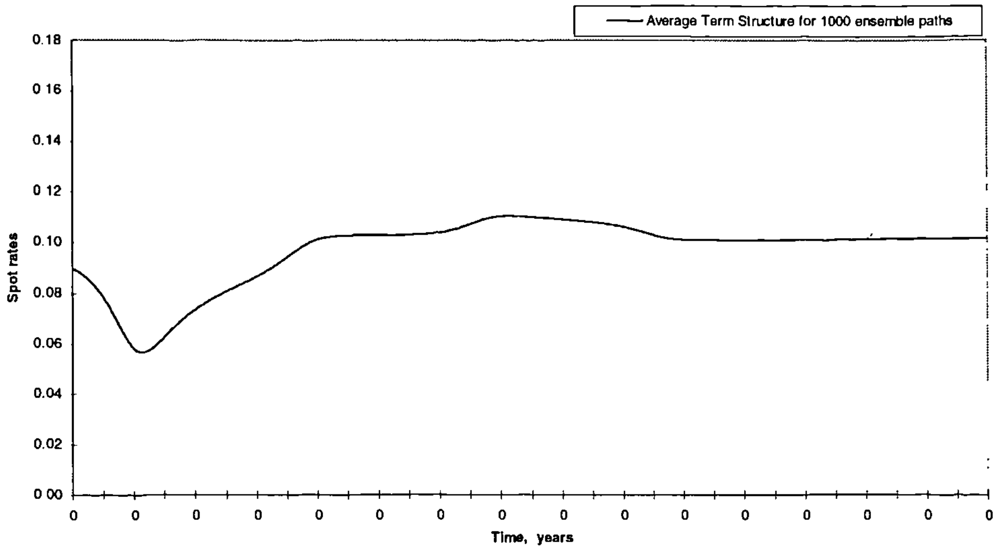


FIGURE 3-17: THE TERM STRUCTURE OF INTEREST RATES: AN ENSEMBLE ESTIMATE

The resemblance between Figure 3-8 and Figure 3-17 is not accidental. Prices calculated using an ensemble approach implicitly assume that there is uncertainty about the value of x. Let us suppose that the market perceives x with noise, so that the perceived value of x is

$$x_t + w_t, \tag{3.4-6}$$

where w_t represents the noise in x_t. (3.4-6) may be interpreted as the market's constant revision of its opinion of the value of x_t as news arrives. The process for r becomes

$$dr = \alpha(x + w - r)dt$$

and r reverts deterministically to $x + w$. It is not necessary to presume that $x + w$ is a continuous process. For instance suppose that w is 'white noise'. Without attempting to rigorously define w one can suppose it has the property that $wdt = \sigma dz$ for some Wiener process z. Hence the process for r becomes

$$dr = \alpha(x - r)dt + \alpha\sigma dz$$

Error in the perception of x is observed as noise in the process for r. Furthermore, since α is large, any error in the perception of x has a disproportionate effect on the noise in r. It might thus be regarded that noise in r is attributable to 'observation error' in x. Uncertainty in the perception of x leads to noise in the process for r, and hence r will acquire a price of risk.

3.4.2 The exercise of control

Control of chaotic systems has been discussed in various contexts, for instance Holyst et al. (1996), Barrett (1996), and the review article by Abarbanel et al. (1993). In this model the behaviour of orbits on the attractor suggests possibilities for control by the monetary authorities. If by fiat of the monetary authorities the variable p could be controlled, and its value varied in response to changes in x and r, relatively small adjustments to the value of p, at the correct point in the cycle, could effectively control interest rates. For instance, the system could be nudged closer to the fixed point at the centre of one of the wings of the attractor. To achieve this it would not be necessary to completely determine p, but only to moderate or increase its rate of movement, in effect by controlling the parameters γ, δ or ϕ. Directly altering the value of p would entail the ability to control either b, k, or u, although if b and u are constrained to be positive, control of these variables alone might be insufficient to achieve the desired effect. Control of k, the variable representing the transactions demand for money, which has been re-interpreted as the availability of transaction credit, might be achieved through credit controls. By regulating the market for transaction credit, the monetary authorities could exert control over k, and hence p.

If direct control of p were possible then another strategy might be effected when r and x have approximately similar values in mid range, in the region 'A' in Figure 3-18. There p is steadily increasing, while r and x remain relatively stable, temporarily. When p is sufficiently large, x and r 'take off' and continue around the attractor. If the value of p could be prevented from increasing, then x and r might be prevented from leaving the mid-range region. For instance if the value of p were to be held at 0 then r would fluctuate about a value that converged to μ.

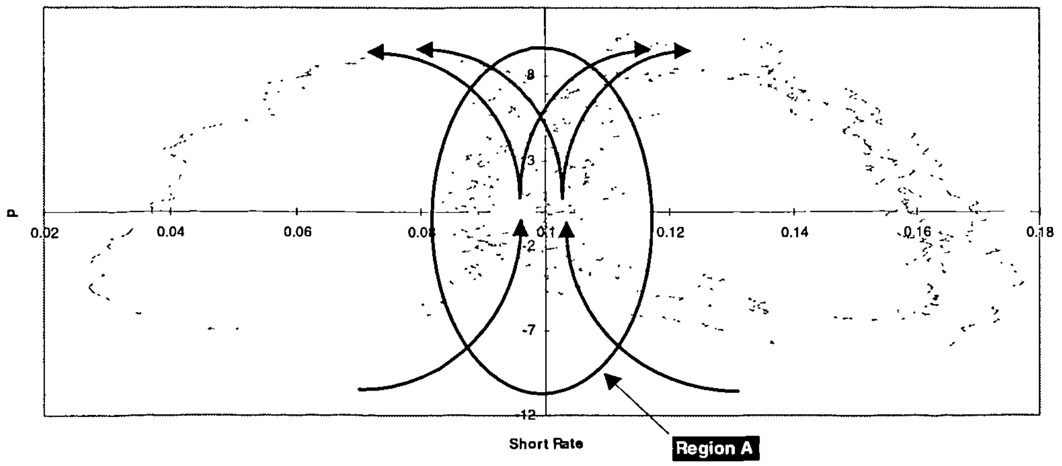


FIGURE 3-18: CONTROL OF p

Yet another form of control might be possible. The variable β represents the speed of adjustment towards equilibrium in the goods market. If it is possible to alter β while keeping α, γ and δ fixed then it may be possible to cause the system to enter non-chaotic regions. The critical value, $\delta_H(\beta)$, at which the additional fixed points become unstable is a function of β. For the illustrative values of α, β and γ it has been seen that

$$\delta = 23 > \delta_H(\beta) \approx 16.94$$

If β were reduced to 0.25 then the critical value of δ_H is raised to approximately 28.46, so that chaotic behaviour is no longer possible with δ = 23.[21,22] β appears to operate as a 'dampening down' variable. The faster the rate of price adjustment in the goods market, the lower the critical value of δ_H at which the system allows chaotic

[21] The calculation assumes that α remains constant when β is altered. This is a reasonable approximation when $\alpha_m > \alpha_y$.

[22] It should be noted that periodic orbits may exist, and in any case the system may take a long time to 'settle down' to a stable orbit or into a fixed point.

behaviour. The result is interesting and perhaps surprising, as it goes counter to the intuition that more rapid an adjustment towards equilibrium the more stable the system.

3.5 Conclusions

It has been seen how the two factor model (2.3-5) from chapter two may be extended by allowing the parameter p to vary. The resultant three factor model fits the description of a dynamic mean model as given in chapter 2. The possibilities for the dynamics of the three factor system have been shown to include chaos. Several realistic examples of behaviour of the short rate are given, with the process exhibiting business cycle type behaviour. Estimation and pricing are discussed with particular reference to the chaotic nature of the system. Methods for investigating non-linear systems are employed to give insight into the description of the interest rate process. This provides evidence of in agreement with the belief that an underlying deterministic process generates the large scale fluctuations .

The complex nature of the model allows economic motivation for the fluctuations observed in the short rate process. The economic foundations also open the way to methods of control and stabilisation policy. Chapters 4 and 5 detail methods for estimating the parameters of the model. The approach in chapter 4 seeks to make use of the inter-relationship between the geometry of the Lorenz system and the economic properties of the model described here.

4. AN ESTIMATION PROCEDURE

4.1 Introduction

Recall from chapter three the system (3.3-4) :

$$dr = \alpha(x - r)dt + \sigma dz$$
$$dx = \beta\big(pr + (1 - p)\mu - x\big)dt$$
$$dp = \gamma\big(\delta - \phi(x - \mu)(r - \mu) - p\big)dt$$

The three factor system is specified by seven parameters and three current values for the state variables. In addition, to use term structure data derived from bond prices, a price of risk must be estimated associated to z. The unknowns fall into four categories :

i) The current values of the state variables, (r_0, x_0, p_0)

ii) The Lorenz parameters, α, γ and δ

iii) Scale parameters, μ, ϕ and β

iv) The volatility and price of risk, σ and λ

The approach given in this chapter seeks to exploit the known geometry of the Lorenz system, as well properties of term structures to investigate the estimation of these parameters. The presence of noise in the system, far from being an unwanted complication, may actually be beneficial; in the approach detailed below this property is exploited. The methods investigated here are heuristic; the topology of the attractor and features of the term structure are used to gain insight into the role of the parameters and approaches to their estimation.

Section 4.2 details a method for uncovering the long term mean, objective drift, price of risk, and volatility of the short rate process. This is approached by exploiting properties of the term structure. Section 4.3 describes methods for uncovering the remaining parameters and unknown state variables of the model. The procedure in this section is based on exploiting discretisations of the three factor system (3.3-4) to provide estimable linear regressions. Section 4.4 applies some of the techniques described in the previous two sections to simulated data to assess their theoretical performance. Section 4.5, drawing on some of the problems of the previous section suggests some alternative methods for estimating individual parameters. This section seeks to reconcile the geometry of the Lorenz system with quantifiable economic values. Section 4.6 concludes, providing some motivation for an improved estimation method in chapter 5.

4.2 Estimation procedure for $\mu, \mu_r(t), \lambda_r, \sigma_r$

The estimation of the parameters $\mu, \mu_r(t), \lambda_r, \sigma_r$ may proceed in a relatively straightforward manner. The parameter μ is allied to the long term mean of the short rate process. Estimation of μ may be done by computing the time average of r_t over a number of cycles.

The volatility and price of risk parameters may be computed from the spatial data by exploiting a well known property of term structures. Given spot rates $r_t(T)$, $T \geq t$, the slope s_t of the term structure at $T = t$ is [1]

$$s_t = \left.\frac{\partial r_t(T)}{\partial T}\right|_{T=t} .$$

It can easily be shown, for instance by expanding $B_t(T)$ in powers of T near $T = t$, that if the short rate follows a diffusion process then[2]

$$s_t = \frac{1}{2}\left(\mu_r(t) - \lambda_r(t)\sigma_r\right) = \frac{1}{2}\tilde{\mu}_r(t) \qquad (4.2\text{-}1)$$

where $\mu_r(t)$, $\lambda_r(t)$ and σ_r are the objective drift, price of risk and volatility of the short rate process at time t. It shall be assumed that $\lambda_r(t) \equiv \lambda_r$ is a constant. In this case

$$\mu_r(t) = \alpha\left(x_t - r_t\right) \qquad (4.2\text{-}2)$$

[1] A method for finding this slope empirically is given in Appendix 4-2

[2] An approximation to the term structure and its derivatives are found Appendix 4-1. This provides the derivation for (4.2-1)

and one may average over a small number of cycles to find that

$$\frac{1}{n}\sum_{t=1}^{n}\mu_r(t)=0$$

hence the value of $\lambda_r\sigma_r$ may be estimated as

$$\lambda_r\sigma_r=-\frac{2}{n}\sum_{t=1}^{n}s_t$$

and the value of $\mu_r(t)$ at each t may be found. Setting

$$\varepsilon_t=\Delta r_t-\mu_r(t)\Delta t \tag{4.2-3}$$

enables the estimation of σ and to identify $\Delta z_t = \varepsilon_t/\sigma$, since by assumption $\varepsilon_t \sim N(0,\sigma^2\Delta t)$.[3] Once σ has been found, the value of λ may also be recovered.

4.3 The remaining parameters: Linear regression approach

The approach of using linear regression to estimate a non-linear/chaotic system may not at first seem tractable. However, knowledge of the geometry of the system and the presence of noise can be exploited to make this approach insightful. Carl Gauss is generally attributed with the discovery of the method of least squares estimation. His first application, to the non-linear estimation problem of the orbit of the planet Ceres, may be thought to have certain parallels to the application here. Whilst this is not an orbit determination problem in the classical sense, the problem is one of estimating the parameters of the system moving around the chaotic attractor. Abarbanel et al (1993) suggest that local linear estimation can be substantially rewarding and is one of the simplest approaches to non-linear modelling. Due to the structure of a chaotic attractor, it can be expected that the orbits will pass through a chosen region repeatedly. Information can then be collected for whole neighbourhoods of phase space, relating how the system evolves from near some $y(n)$ to the whole set of points near $y(n+1)$.

[3] The ε_t are by assumption independent, normal and serially uncorrelated. Tests for these characteristics serve to determine whether aspects of the model are correctly specified.

Examining the local structure of individual neighbourhoods using linear methods can build up a model of the global process, neighbourhood by neighbourhood. The approximations needed for the system (3.3-4) to be estimated in this manner are, however, a compromise for the relatively simplicity of the technique.

4.3.1 *Estimation procedure for* α

α may be estimated by exploiting knowledge of the Lorenz attractor and the presence of noise. In Figure 3-7 when r is at mid values and in the process of switching from one wing of the attractor to another, both r and x are approximately equal to μ for an extended period. It may expected that it is possible to identify a subsequence t_i of indexes such that $\mu_r(t_i) = \alpha(x_{t_i} - r_{t_i}) \approx 0$. Since $\Delta x_t \approx 0$ when $r_t \approx x_t \approx \mu$ it may supposed that

$$\mu_r(t_i + 1) = \alpha(x_{t_i+1} - r_{t_i+1}) \approx \alpha(x_{t_i} - r_{t_i} - \varepsilon_{t_i}) \approx -\alpha\varepsilon_{t_i}$$

Hence the regression

$$\Delta r_{t_i+1} = -\hat{\alpha}\varepsilon_{t_i}\Delta t + \varepsilon_{t_i+1} + \eta_{t_i}$$

or a variant thereof, will enable the identification of α, and hence the sequence of values x_t may be deduced from the known values of $\mu_r(t)$.

4.3.2 *Estimation procedure for* β

Starting from the discretisation for Δx_t :

$$\Delta x_t = p_t\beta(r_t - \mu)\Delta t - \beta(x_t - \mu)\Delta t \tag{4.3-1}$$

and supposing that the path for x_t has been recovered via the estimation of α, a procedure for estimating β may proceed. If the model of interest is the two factor model, then the unknowns in (4.3-1) are the parameters p and β. In this case a regression may be run of the form :

$$\Delta x_t = \hat{\beta}_1(r_t - \mu)\Delta t - \hat{\beta}_2(x_t - \mu)\Delta t + \eta_t \tag{4.3-2}$$

where the coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$ represent estimates of $p\beta$ and β respectively, allowing both p and β to be found. For the two factor model, all of the parameters that specify the system will then have been found. In practice, this estimation may prove problematic due to the high degree of colinearity between $(r_t - \mu)$ and $(x_t - \mu)$. In essence, the problem of colinearity is one of sampling and the consequences of it for the estimates cannot be determined *a priori*.

If the system of interest is the three factor system, then p is a time dependent state variable and unknown. As such, some method has to be found for excluding the term in p_t from the expression (4.3-1) if it is to be the basis for a regression formulation for estimating β. This may entail a procedure whereby a subsequence of indexes t_i is found such that $(r_{t_i} - \mu)$ is close to 0 and $(x_{t_i} - \mu)$ is relatively far from 0. As noted above, it can be expected that r_t and x_t will be close most of the time. Further, the subsequence must be selected such that $(x_{t_i} - \mu)$ is larger than $(r_{t_i} - \mu)$ by at least a factor κ, where κ is an expected average absolute value of p_t. Finding a subsequence that satisfies these criteria may be problematic.

4.3.3 Estimation procedure for α and β concurrently

Given the undesirable nature of estimating parameters recursively and the problems noted in section 4.3.2 for the estimation of β, a procedure for estimating α and β concurrently is postulated. Starting from the objective drift for the process for r:

$$\begin{aligned}
\mu_r(t+1) &= \alpha(x_{t+1} - r_{t+1}) \\
&= \alpha(x_t - r_t + \Delta x_t - \Delta r_t) \\
&= \alpha(x_t - r_t) + \alpha(\Delta x_t) - \alpha(\Delta r_t) \\
&= \mu_r(t) + \alpha(\Delta x_t) - \alpha\mu_r(t)\Delta t - \alpha\varepsilon_t
\end{aligned} \tag{4.3-3}$$

The unknown quantities in (4.3-3) are Δx_t and the parameter α (assuming that $\mu_r(t)$ and ε_t have been found via the procedure in section 4.2 for instance). Since Δx_t cannot be constrained directly, it is desirable to remove it via a reformulation.

The expression for Δx_t is given by (4.3-1) above and the objective drift of r is given by :

$$\mu_r(t) = \alpha(x_t - r_t).$$

Solving the expression $\mu_r(t)$ for x_t gives

$$x_t = \frac{\mu_r(t)}{\alpha} + r_t \qquad \text{(for alpha not close to 0)}$$

Substituting this expression for x_t into Δx_t :

$$\begin{aligned}\Delta x_t &= p_t\beta(r_t - \mu)\Delta t - \beta\left(\frac{\mu_r(t)}{\alpha} + r_t - \mu\right)\Delta t \\ &= p_t\beta(r_t - \mu)\Delta t - \beta(r_t - \mu)\Delta t - \beta\left(\frac{\mu_r(t)}{\alpha}\right)\Delta t\end{aligned}$$

Substitute for Δx_t back into (4.3-3)

$$\begin{aligned}\mu_r(t+1) = \mu_r(t) &+ \alpha\left[p_t\beta(r_t - \mu)\Delta t - \beta(r_t - \mu)\Delta t - \beta\left(\frac{\mu_r(t)}{\alpha}\right)\Delta t\right] - \\ &- \alpha\mu_r(t)\Delta t - \alpha\varepsilon_t \\ &= \mu_r(t) + \alpha\left[(p_t - 1)\beta(r_t - \mu)\Delta t - \beta\left(\frac{\mu_r(t)}{\alpha}\right)\Delta t\right] - \alpha\mu_r(t)\Delta t - \alpha\varepsilon_t \\ &= \mu_r(t) + \alpha\beta(p_t - 1)(r_t - \mu)\Delta t - \beta\mu_t(t)\Delta t - \alpha\mu_r(t)\Delta t - \alpha\varepsilon_t\end{aligned}$$

from which an expression for $\Delta\mu_r(t)$ can be found :

$$\begin{aligned}\Delta\mu_r(t) &= \mu_r(t+1) - \mu_r(t) \\ &= \alpha\beta(p_t - 1)(r_t - \mu)\Delta t - (\alpha + \beta)\mu_r(t)\Delta t - \alpha\varepsilon_t\end{aligned} \qquad (4.3\text{-}4)$$

Now, the unknowns in (4.3-4) are α and β (the parameters) and p_t. Again, if the model of interest is the two factor model then p will be a time invariant parameter. In this case, a regression can be estimated of the form :

$$\Delta\mu_r(t) = \hat{\beta}_1(r_t - \mu)\Delta t + \hat{\beta}_2\mu_r(t)\Delta t + \hat{\beta}_3\varepsilon_t + \eta_t$$

where

$$\beta_1 = \alpha\beta(p-1)$$
$$\beta_2 = -(\alpha+\beta)$$
$$\beta_3 = -\alpha$$

allowing α, β and p to be estimated.

For the three factor model, p is the third unknown state variable and some method for removing it from the expression for $\Delta\mu_r(t)$ will allow an estimation to be performed. If we can select a subsequence of time indexes such that the term in p_t is 0 (relative to $\mu_r(t)\Delta t$ and ε_t) then one can estimate a regression of the form:

$$\Delta\mu_r(t) = \hat{\beta}_1\mu_r(t)\Delta t + \hat{\beta}_2\varepsilon_t + \eta_t \qquad (4.3\text{-}5)$$

where

$$\hat{\beta}_1 = -(\alpha+\beta)$$
$$\hat{\beta}_2 = -\alpha$$

What criteria should be used to select the subsequence? A restriction can be placed upon $(r_t-\mu)$ such that $\alpha\beta(p_t-1)(r_t-\mu)\Delta t$ is 'close' to 0. This term will essentially represent the error in the regression. To ensure that the error is small compared to the explanatory variables, one can concurrently select a subsequence where x_t is far from r_t, by choosing time indexes with $\mu_r(t)$ large.

4.3.4 *Recovering the path for p*

In the three factor model, the path for the state variable p has to be recovered. Once estimates of x_t and β have been found, this may be done by solving the expression (4.3-1) for p_t.

$$\Delta x_t = p_t\beta(r_t-\mu)\Delta t - \beta(x_t-\mu)\Delta t$$
$$p_t = \frac{\Delta x_t + \beta(x_t-\mu)\Delta t}{\beta(r_t-\mu)\Delta t} \qquad (4.3\text{-}6)$$

In practice, this may be problematic when r_t is close to μ, leading to very noisy estimates for p_t.

4.3.5 Estimation procedure for γ, δ, and ϕ

Under the supposition that the path for p_t has been recovered, a procedure for estimating the parameters γ, δ, and ϕ may proceed. Utilising the discretisation of the third in the system of differential equations (3.3-4):

$$\begin{aligned} \Delta p_t &= \gamma(\delta - \phi(r_t - \mu)(x_t - \mu) - p_t)\Delta t \\ &= \gamma\delta\Delta t - \gamma\phi(r_t - \mu)(x_t - \mu)\Delta t - \gamma p_t \Delta t \end{aligned}$$

A regression of the following form can be run:

$$\Delta p_t = \hat{\beta}_0 - \hat{\beta}_1(r_t - \mu)(x_t - \mu) - \hat{\beta}_2 p_t \Delta t + \xi_t \,. \tag{4.3-7}$$

Then the parameter estimates for γ, δ and ϕ can be found from:

$$\begin{aligned} \beta_2 &= -\gamma & & & \gamma &= -\beta_2 \\ \beta_1 &= -\gamma\phi\Delta t & &\Rightarrow & \phi &= \frac{\beta_1}{\beta_2 \Delta t} \\ \beta_0 &= \gamma\delta\Delta t & & & \delta &= \frac{-\beta_0}{\beta_2 \Delta t} \end{aligned}$$

4.4 Application to simulated data

This section seeks to quantify the applicability of the estimation procedures in sections 4.2 and 4.3. The methods outlined above are applied to simulated data to assess their ability to recover the paths for the hidden state variables x and p, and the parameters that identify the system. One of the significant drawbacks in the application of the approaches here, is in the use of parameter estimates recursively. For instance, the estimates of the state variable p are found based upon estimates of α, β, and μ. The estimates of the parameters γ, δ, and ϕ are then made conditional on the path for p.

A second disadvantage of the methods in section 4.3, is that they are based upon a first order discretisation of the system (3.3-4). Use of second order discretisations is not always advantageous. An investigation surrounding this is made below.

In the simulated results as presented below, parameter values similar to those used in chapter 3 are employed. These are tabulated below.

TABLE 4-1: PARAMETER VALUES USED FOR SIMULATED ANALYSIS

Lorenz parameters		Scale parameters		Volatility	
α	5	μ	0.1	σ	0.025
γ	5/12	ϕ	43456.79		
δ	23	β	0.5		

Simulated sample paths are found by numerically integrating the system (3.3-4) as in chapter 3. For this analysis, ten years of weekly data is simulated, consisting of 520 data points; $\Delta t=1/52$. Monte Carlo procedures used incorporate the antithetic variance reduction technique.

Estimation of σ and μ from simulated data is trivial. As noted in section 4.2 an estimate of μ may be found from the time average of r_t over a number of cycles. Assuming the objective drift for the process for r has been found, σ may be uncovered via the distribution of ε_t given by (4.2-3). Monte Carlo results for these parameters are given in Table 4-2 below. Two hundred sample paths are used in the estimation.

TABLE 4-2: MONTE CARLO ESTIMATES OF μ AND σ

	μ	σ
Estimate	0.1015	0.02381
S.E.	*0.0006*	*0.00005*

The values in Table 4-2 may be compared with those in Table 4-1. The estimate for the long term mean is over the benchmark value by some 2.6 times the standard error. Given that the Lorenz system has a natural cyclical pattern, it is likely that the average number of cycles implied by the parameter values and time step used here correspond to a fractional value. Only if the parameter values, starting values and time scale are tuned deterministically to give a time average of the short rate equal to μ could it be expected that the Monte Carlo estimate would be accurate. Here the values used to simulate the system are chosen for their heuristic ability to qualitatively describe the short rate process. Empirically, it is judicial to ensure that the number of time indexes chosen to estimate μ corresponds to as complete a number of cycles as possible.

The point estimate of σ is also somewhat divergent from its theoretical value. It is the case here that the objective drift of the process for r in (4.2-2) has been obtained from a first order discretisation. A full implementation of the method in section 4.2 would produce Monte Carlo simulations of the term structure, as in Figure 3-8, to estimate the objective drift, price of risk and volatility for the short rate.

4.4.1 Estimating α and β

The method as discussed in section 4.3.3 is applied to simulated data using a regression of the form (4.3-5). To enable (4.3-5) to be applied, some restriction has to be placed upon $(r_t - \mu)$. This is necessary for (4.3-5) to be a good approximation to the expression (4.3-4).

It is required to choose subsequences of time indexes such that $(p_t - 1)(r_t - \mu)$ is small, and does not contribute a significant amount of information in the construction of $\Delta\mu_r(t)$. Choosing the form of the restriction depends partly on one's expectation of the parameters α, and β and the range of values over which p_t will vary. It is required that the term $\alpha\beta(p_t - 1)(r_t - \mu)\Delta t$ is small relative to $(\alpha + \beta)\mu_r(t)\Delta t$ and $\alpha\varepsilon_t$. If α and β are significantly less than 0 then the factor $\alpha\beta$ will be significantly less than (α + β). Conversely, if α and β are significantly greater than 0 then the factor $\alpha\beta$ will be significantly greater than (α + β). For the heuristic values used in the simulations here, and given in Table 4-1, it is the case that $\alpha\beta$ = 2.5 and (α + β) = 5.5. With an average absolute value for p_t of 5, the magnitudes of the terms in the expression (4.3-5) can be assessed.

$$\begin{aligned} |\alpha\beta(p_t - 1)(r_t - \mu)\Delta t| &\approx |10(r_t - \mu)\Delta t| \\ (\alpha+\beta)\mu_r(t)\Delta t &= 5.5\mu_r(t)\Delta t \\ |\alpha\varepsilon_t| = |\alpha\sigma\Delta z_t| &\approx |0.125\sqrt{\Delta t}| \end{aligned} \quad (4.4\text{-}1)$$

In reality, one is at liberty to place restrictions on both $(r_t - \mu)$ and $\mu_r(t)$; if a restriction is placed for $|(r_t - \mu)| < \kappa$, then based upon (4.4-1) a second restriction may be placed for $|\mu_r(t)| > 2\kappa$. With the value chosen for Δt here of 1/52 an average absolute value of $\alpha\varepsilon_t$ will be approximately 1/56 or 0.017. Setting κ to a value of 0.01 or less should ensure that (4.3-5) is a good approximation to (4.3-4). An average absolute maximal value of $\alpha\beta(p_t - 1)(r_t - \mu)\Delta t$ in the case of $\kappa = 0.01$ would be 0.0019. As κ is reduced towards 0, the approximation should perform better. The trade-off then comes in that fewer observations will be available with larger implied standard errors for the parameter estimates.

To quantify this approach a Monte Carlo procedure is run to gain estimates of the parameters α and β under the scheme as detailed above, for different values of the constraint value κ. Two hundred sample paths are used for each estimation. The results are tabulated in Table 4-3 below.

TABLE 4-3: MONTE CARLO SIMULATION ESTIMATES FOR α AND β[4]

Constraints $\lvert(r_t-\mu)\rvert<\kappa$ $\lvert\mu_r(t)\rvert>2\kappa$	$\hat{\alpha}$	$\hat{\beta}$	No. of observations
$\kappa = 0.01$	4.915 *0.003*	0.143 *0.027*	107.920 *2.385*
$\kappa = 0.005$	4.918 *0.003*	0.386 *0.017*	74.075 *1.997*
$\kappa = 0.0025$	4.919 *0.003*	0.456 *0.013*	42.440 *1.296*
$\kappa = 0.00125$	4.919 *0.003*	0.484 *0.014*	22.400 *0.743*

From the Table 4-3 it can be seen that the estimates for α and β improve as the value of κ is increased. This is in line with expectations, based upon the discussion above. As the value of the constraint increases, the number of observations available decreases. For the most restrictive value of κ, the standard error on the parameter β is marginally larger than for $\kappa = 0.0025$. It is noticeable that the estimates are biased downward from their benchmark values, as in Table 4-1. This can be explained as a symptom of the first order discretisation which has been used to find the reduced form expression (4.3-4). The use of the discretisation (4.3-4) can be thought of as an approximation to the exact discretisation, where the regressors are proxies for the real regressors and measured with error. To fully understand the effect of the discretisation, some quantification has to be made of the error introduced, specifically, the distribution of the error. In the case where the measurement errors are independent Gaussian, the effects are well document for the linear regression case. See for instance Carroll, Ruppert and Stefanski (1995).

If, instead of observing the real regressor X, some proxy is observed $W = X + U$, where U is independent of X and distributed with mean zero and variance σ_U^2. The least

[4] Standard errors are given in italics under the estimates.

squares regression of Y on W will not yield a consistent estimate of β_X; the estimate will be biased by some factor λ, such that the observed estimate will be $\lambda\beta_X$. The factor λ is given by :

$$\lambda = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_U^2} < 1$$

The use of W as a proxy for X in the regression against Y will yield an estimator that is biased toward 0.

The distributions of the discretisation errors are not evaluated here, and, in all likelihood they will not be independent of the regressors. However, the above discussion serves to show how a negative bias in the parameter values may be accommodated.

A logical step to consider for elimination of the bias introduced with the discretisation, is the use of a second order approximation to the system (3.3-4). Unfortunately, at this stage in the estimation process this is not useful. The second order approximation to the expression introduces terms in $\Delta r_t p_t$. In order to remove these, it would necessitate an extra restriction for Δr_t to be close to 0. Recall,

$$\Delta r_t = \mu_r(t)\Delta t + \varepsilon_t$$

and the form of the reduced form expression for $\Delta\mu_r(t)$ which is the basis for the estimation of α and β:

$$\Delta\mu_r(t) = \alpha\beta(p_t - 1)(r_t - \mu)\Delta t - (\alpha + \beta)\mu_r(t)\Delta t - \alpha\varepsilon_t$$

Restricting Δr_t and $(r_t - \mu)$ to be close to 0 concurrently makes the use of the above expression impractical for estimating α and β.

4.4.2 Recovering the path for p

The path for p_t may be found via the expression (4.3-6), reproduced below.

$$p_t = \frac{\Delta x_t + \beta(x_t - \mu)\Delta t}{\beta(r_t - \mu)\Delta t}$$

However, this estimate will be noisy when (r_t - μ) is close to 0. Figure 4-1 below shows how the noise may be manifested in a reconstructed path for p_t using the expression (4.3-6). A Monte-Carlo sample path is produced using the values in Table 4-1 and the path for p_t recovered. The exact values for the quantities on the right hand side of (4.3-6) are used.

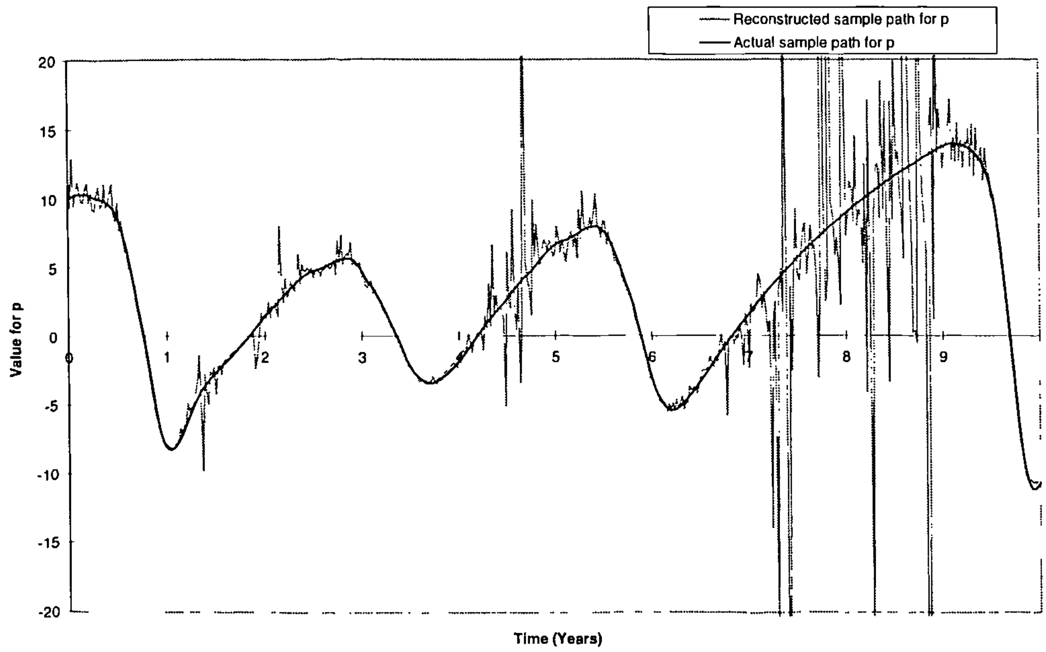


FIGURE 4-1: RECONSTRUCTED SAMPLE PATH FOR p_t

It is clear from Figure 4-1 that the recovered sample path for p_t is very poor. On closer inspection, it can be seen that the recovered sample path is relatively accurate over the region where p_t is decreasing. When p_t is increasing the accuracy deteriorates markedly. The supposition is, from the form of (4.3-6), that when r_t is close to μ the estimates of p_t are poor. Looking at the discretisation of the differential equation for p,

$$\Delta p_t = \gamma\delta\Delta t - \gamma\phi(r_t - \mu)(x_t - \mu)\Delta t - \gamma p_t \Delta t$$

it can be seen that when r_t is close to μ, p_t is reverting to its long term rate, δ. Here, the value of δ is set at 23, a value which p_t never attains such that when r_t is close to μ, p_t is rising. As discussed in section 4.3.5 above, the recovered path for p_t allows for an estimation procedure for the remaining parameters to proceed. Some method of recovering, at least partial, quality estimates for the series of p_t will prove beneficial to the estimation procedure.

Estimates of p_t in Figure 4-1 are gained from solving the discretised version of the differential equation for x_t, leading to the expression (4.3-6). Utilising a second order approximation to the differential equation for x_t which is then solved for p_t may provide superior results. The resultant expression is then :

$$p_t = \frac{\frac{2\Delta x_{t-1}}{\beta \Delta t} + x_{t-1} + x_t - 2\mu - p_{t-1}(r_t - \mu)}{r_{t-1} - \mu}$$

with p_t dependent upon the previous period value and scaled by $(r_{t-1} - \mu)$. Again estimates of p_t approach a point of singularity as r_t becomes close to μ. In practice, estimates of p_t found by this approach are worse then those found by (4.3-6), not least, in part due to the propagated error from one time period to the next.

Another approach may be taken based upon the observation that the estimates of p_t gained from (4.3-6) are relatively accurate over the range where r_t is far from μ. As such, subsequences of time indexes may be selected for $|(r_t - \mu)| > \kappa$, where as κ increases the cut off accuracy of the chosen time indexes will increase.

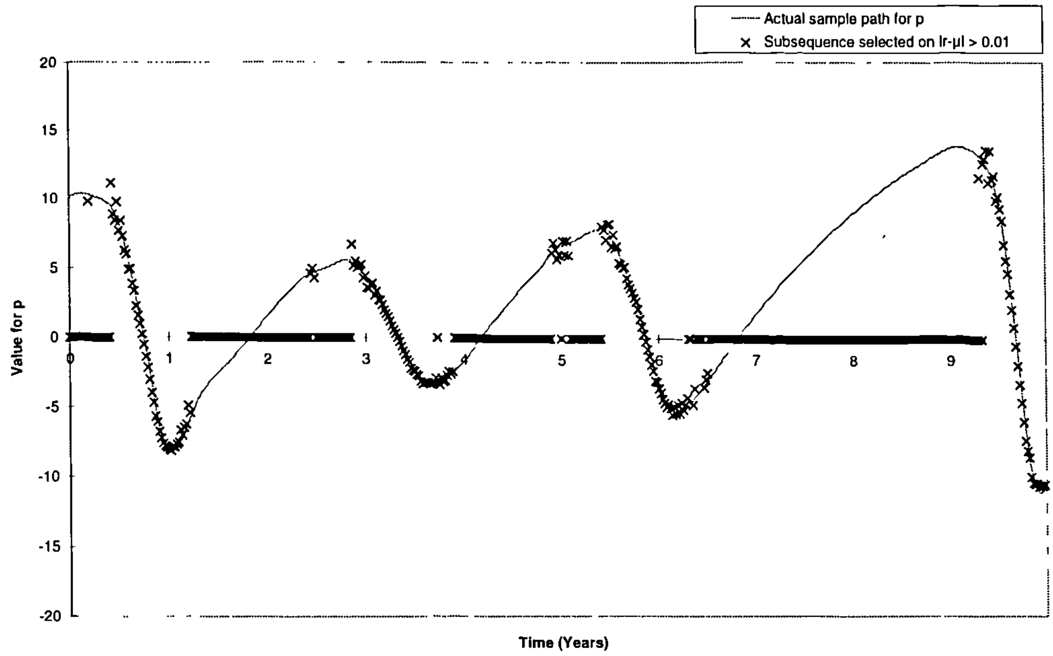


FIGURE 4-2: SUBSEQUENCE OF TIME INDEXES OF RECOVERED SAMPLE PATH FOR p_t CHOSEN UNDER CRITERION $|r_t - \mu| > 0.01$

Figure 4-2 above shows how the restriction chooses points along the recovered sample path. The subsequence of time indexes chosen, consists mainly of those where the change in p_t is negative, and the accuracy of the estimates is greater than those estimates rejected.

4.4.3 Estimating γ, δ, and ϕ

Estimation of γ, δ, and ϕ may be effected via the regression (4.3-7). In practice, this may be problematic due to the poor nature of the recovered series for p_t as discussed above. Here, to investigate the use of (4.3-7) Monte Carlo procedures are run to gain estimates of γ, δ, and ϕ. Two hundred sample paths are simulated to find the Monte-Carlo estimates and the associated standard errors.

To assess how well (4.3-7) performs in its ability to estimate γ, δ, and ϕ, the regression is first run using the exact simulated values for p_t as opposed to the recovered estimates from (4.3-6). No restrictions are placed upon the time indexes used for the regression. Results from applying this estimation procedure are shown in Table 4-4 below.

TABLE 4-4: MONTE CARLO ESTIMATES FOR γ, δ, ϕ USING EXACT p_t

	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\phi}$
Estimate	0.4897	19.8011	36844.18
S.E.	*0.0006*	*0.0277*	*48.45*

From Table 4-4 it can be seen that the estimates of γ, δ, and ϕ are biased against their theoretical benchmark values as in Table 4-1. In all cases, the estimated values are more than 100 standard deviations away from the true value. Again, the biases involved here are mainly due to the first order discretisation employed in the construction of the regression.

An equivalent regression may be estimated based upon a second order discretisation. Taking a Taylor series expansion for $p(t)$ up to terms of order two:

$$p(t+\Delta t) = \underbrace{p(t) + p'(t)\cdot\Delta t}_{\text{Simple Euler}} + p''(t)\cdot\frac{\Delta t^2}{2} + \ldots$$

$$\Delta p(t) \approx p'(t)\cdot\Delta t + p''(t)\cdot\frac{\Delta t^2}{2}$$

$$\left\{\begin{array}{l} p'(t) = \gamma\left[\delta - \phi(x(t)-\mu)(r(t)-\mu) - p(t)\right] \\ p''(t) = \gamma\left[-\phi(x'(t)\cdot r(t)) - \phi(r'(t)\cdot x(t)) + \phi\mu\cdot x'(t) + \phi\mu\cdot r'(t) - p'(t)\right] \end{array}\right. \Rightarrow$$

$$\begin{aligned}\Delta p(t) \approx & \left[\gamma\left(\delta - \phi(x(t)-\mu)(r(t)-\mu) - p(t)\right)\right]\Delta t + \\ & + \left[\gamma\left(-\phi\frac{\Delta x(t)}{\Delta t}\cdot r(t) - \phi\frac{\Delta r(t)}{\Delta t}\cdot x(t) + \phi\mu\frac{\Delta x(t)}{\Delta t} + \phi\mu\cdot\frac{\Delta r(t)}{\Delta t} - \frac{\Delta p(t)}{\Delta t}\right)\right]\Delta t\end{aligned}$$

Collecting terms yields the expression

$$\begin{aligned}\Delta p(t) \approx & \frac{\gamma}{\left(\frac{1}{\Delta t} + \frac{\gamma}{2}\right)}\cdot\left\{\delta - \phi\left[(x(t)-\mu)(r(t)-\mu) - \frac{1}{2}(\mu - r(t))\Delta x(t)\right.\right. \\ & \left.\left. - \frac{1}{2}(\mu - x(t))\Delta r(t)\right] - p(t)\right\}\end{aligned}$$

from which a regression can be estimated of the form :

$$\Delta p(t) = \hat{\beta}_0 + \hat{\beta}_1\left[(x(t)-\mu)(r(t)-\mu) - \frac{1}{2}(\mu - r(t))\Delta x(t) - \frac{1}{2}(\mu - x(t))\Delta r(t)\right] +$$
$$+\hat{\beta}_2 p(t)\Delta t + \xi_t.$$

(4.4-2)

where

$$\beta_2 = \frac{-\gamma}{\left(1+\frac{\gamma\Delta t}{2}\right)} \qquad \gamma = \frac{-2\beta_2}{(2+\Delta t\beta_2)}$$

$$\beta_1 = \frac{-\gamma\phi}{\left(\frac{1}{\Delta t}+\frac{\gamma}{2}\right)} \qquad \Rightarrow \qquad \phi = \frac{\beta_1}{\beta_2\Delta t}$$

$$\beta_0 = \frac{\gamma\delta}{\left(\frac{1}{\Delta t}+\frac{\gamma}{2}\right)} \qquad \delta = \frac{-\beta_0}{\beta_2\Delta t}$$

Running a Monte-Carlo estimation procedure using (4.4-2) produces the estimates reported in Table 4-5 below. As in the estimation for the first order discretisation above, the exact values of p_t are used and no restriction is placed on the time indexes used.

TABLE 4-5: MONTE CARLO ESTIMATES FOR γ, δ, ϕ USING EXACT p_t AND SECOND ORDER TERMS

	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\phi}$
Estimate	0.41520	23.089	43704.53
S.E.	*0.00003*	*0.002*	*3.73*

The results as presented in Table 4-5 show improved point estimates over those found using only a first order discretisation. The approach used to generate the estimates in Table 4-5 is, however, not applicable in reality, as the series for p_t will be recovered via (4.3-6). As seen in section 4.4.2, subsequences can be selected via a criterion on $(r_t - \mu)$ such that the chosen time indexes are those with better than average fit to the actual path for p_t.

The regression (4.4-2) is applied using the recovered series for p_t with subsequences selected for $(r_t - \mu) > \kappa$. It is expected that as κ increases, the estimates should approach those in Table 4-5.

TABLE 4-6: MONTE CARLO ESTIMATES FOR γ, δ, ϕ USING RECOVERED p_t AND SELECTED TIME INDEXES

Constraint $\lvert(r_t - \mu)\rvert > \kappa$	$\hat{\gamma}$	$\hat{\delta}$	$\hat{\phi}$	No. of observations
$\kappa = 0.01$	1.10317	11.23	17848.75	339.62
	0.02393	*1.33*	*2063.62*	*3.69*
$\kappa = 0.02$	0.82029	17.66	28751.94	206.76
	0.01876	*1.33*	*1795.09*	*2.53*
$\kappa = 0.03$	0.86715	14.60	25968.78	119.92
	0.02347	*0.91*	*1610.64*	*1.12*
$\kappa = 0.04$	1.06237	10.43	20079.21	59.30
	0.03742	*1.81*	*2123.05*	*1.06*
$\kappa = 0.05$	2.50582	-1.10	8141.57	21.32
	0.55660	*1.58*	*1911.18*	*0.64*

The estimation that finds the best point estimates for γ, δ and ϕ, is that for $\kappa = 0.02$. This also yields the lowest standard error for the estimate of γ. For $\kappa = 0.05$, the estimates are, unexpectedly, much poorer than those for $\kappa = 0.02$. In the case of $\kappa = 0.05$, the estimate of δ is negative. Indeed, all of the estimates in Table 4-6 are poor compared with the both the theoretical benchmark values of Table 4-1 and the expected Monte Carlo estimates gained using the actual series for p_t given in Table 4-5. The implication is that this approach would not be reliable in estimating the parameters from empirical data. Tentative results using longer series and daily data were not found to alter the estimates significantly.

Further methods may be sought for recovering estimates of γ, δ and ϕ based on the approach described in this section. Given that with perfect estimates for p_t application of the regression (4.4-2) is viable, some method for removing the noise of the recovered series for p_t may be beneficial. Notably, a smoothing technique may be applied to the data to remove the undue noise. A lowpass filter applied to the recovered series for p_t would enable the large scale dynamics of the system to be uncovered.· Given that the series for p_t is smooth by virtue of having no noise directly associated with it, this approach may be particularly useful. Lowpass filters smooth by computing a positively weighted average across an interval. Furthermore, lowpass filters are generally more effective then a simple moving average filter, the effect of which is irregular on the cyclical components of the series. Many variants of lowpass filters exist. See, for example, Masters (1995). One type of filter that may be particularly useful in this instance is the Savitzky Golay filter, which is applicable to series that vary smoothly and at a relatively slow rate.

The Savitzky Golay filter works in a different manner to most lowpass filters. To compute the filtered value of a point it collects the values of the points in the surrounding neighbourhood. A polynomial of degree between two and six is fitted to the data over the region. The polynomial evaluated at the central point yields the filtered value. The filter can be made wide to capture a large amount of information, with a polynomial degree high enough to follow the variations in that interval. From looking at the sample path for the actual series for p_t in Figure 4-1, it is viable to think that a low order polynomial could describe the series over a relatively long time interval. For further details and implementation see Press et al (1992)

The results of applying such a filtering technique are not investigated here. Preliminary results from the application of moving average and lowpass filters to the recovered series for p_t did not prove encouraging. Application of the Savitzky Golay filter may prove more rewarding due to its characteristics noted here.

4.5 Some alternative approaches

In the preceding sections of this chapter, it has been seen how an estimation approach may proceed. Given the piecemeal nature of the approach, there are several other properties that may be exploited for estimating individual parameters. There now follows a discussion of some of the properties of the system that may be exploited in estimating the parameters. In particular, the methods described here may provide a means for uncovering the values of γ, δ and ϕ which have proved difficult to estimate in section 4.4.3.

The speed of the system is determined by the three reversion parameters, α, β and γ. If two of these have been estimated with certainty, then the third may be found by matching the speed of the system to empirical data. The relative sizes of these parameters are related directly to the scale of the time axis. By altering the parameter β, while keeping α/β and γ/β fixed the system is effectively speeded up or slowed down. In this way the scale over which fluctuations occur from high rates to low rates can be controlled. This can then be matched to the empirical business cycle.

The two unstable fixed points of the model around which the attractor forms the two lobes are found at :

$$r = x = \mu \pm \sqrt{\frac{\delta - 1}{\phi}}$$

It may be possible to estimate the location of the fixed points of the system by observing the empirical movement of the short rate over a number of cycles. This will then allow the uncovering of the scale parameter ϕ as a function of μ and δ. As discussed above μ can be found as the time average of the short rate over a number of cycles. The state variable p represents the availability of transactions credit in the economy. From (3.2-3), the long run rate for p, denoted $\bar{p}$, is given by the expression :

$$\bar{p} = \delta - \phi(x - \mu)(r - \mu)$$

It may be possible to determine the value of $\bar{p}$ from economic fundamentals. If both the long run rate for p and the location of the fixed points of the attractor can be estimated, then both δ and ϕ can be estimated.

4.6 Conclusion

This chapter has investigated a variety of methods that may be used to estimate the parameters of the model and uncover the hidden state variables. The major drawback of the approach is that parameter estimates are made sequentially; later parameter estimates are made conditional on previous parameter estimates. This is undesirable in that errors in the estimation of one parameter will be propagated in the estimation of subsequent parameters. Monte Carlo simulations of the estimation procedures for μ, σ, α and β performed well. It would be expected that empirically these parameters can be recovered satisfactorily. Estimates of the path for the state variable p_t were found to be very noisy and liable to yield poor estimates for γ, δ, and ϕ. Alternative methods based on exploiting the geometry of the Lorenz system are mooted.

Application to empirical data could proceed based upon the methods detailed here. However, the problems observed make it undesirable to do so. Further revisions of the procedures, as discussed, may make application more tenable. Chapter 5 seeks to address some of the inadequacies in the estimation process described here. Notably, it is sought to provide a compact estimation process for estimating all the parameters of the model concurrently. Additionally, the procedures in chapter 5 seek to make more extensive use of spatial data.

APPENDIX 4-1: BOND PRICING EQUATION APPROXIMATION

From (3.4-1), the bond pricing equation is given by

$$rB(\tau)=\frac{1}{2}\sum_{i,j=r,x,p}\rho_{ij}\sigma_i\sigma_j\frac{\partial^2 B(\tau)}{\partial i\partial j}+\sum_{i=r,x,p}\tilde{\mu}_i\frac{\partial B(\tau)}{\partial i}+\frac{\partial B(\tau)}{\partial t}, \quad \text{(A4-1)}$$

where $\tau = T - t$ and $\tilde{\mu}_i = (\mu_i - \lambda_i\sigma_i)$. Expanding $B(\tau)$ in powers of τ gives

$$B(\tau)=1+a\tau+b\tau^2+c\tau^3+\ldots$$

Then matching powers of τ on each side of the bond pricing equation (A4-1) allows the coefficients $a,\ b,\ c,$ to be uncovered :

$$\tau^0: \quad -a=r \quad \Rightarrow \quad a=-r$$

$$\tau^1: \quad -2b-\tilde{\mu}_r=-r^2 \quad \Rightarrow \quad b=\frac{r^2-\tilde{\mu}_r}{2}$$

$$\tau^2: \quad -3c+r\tilde{\mu}_r-\frac{1}{2}\sum_{i=r,x,p}\tilde{\mu}_i\frac{\partial\tilde{\mu}_r}{\partial i}+\frac{1}{2}\left[\sigma_r^2-\frac{1}{2}\sum_{i,j=r,x,p}\rho_{ij}\sigma_i\sigma_j\frac{\partial\tilde{\mu}_r}{\partial i\partial j}\right]=r\cdot b$$

$$\Rightarrow \quad 2c=r\tilde{\mu}_r-\frac{1}{3}\left[\sum_{i=r,x,p}\tilde{\mu}_i\frac{\partial\tilde{\mu}_r}{\partial i}-\sigma_r^2+\frac{1}{2}\sum_{i,j=r,x,p}\rho_{ij}\sigma_i\sigma_j\frac{\partial\tilde{\mu}_r}{\partial i\partial j}+r^3\right]$$

Constructing the logarithmic series for the bond pricing equation expansion allows the expression for the term structure of interest rates to be evaluated.

$$\begin{aligned}R(t,T)&=-\frac{1}{\tau}\ln(B(\tau))\\&=-\frac{1}{\tau}\ln\left(1+\left(a\tau+b\tau^2+c\tau^3+\ldots\right)\right)\\&\approx-\left(a+b\tau+c\tau^2-\frac{1}{2}a^2\tau-ab\tau^2+\frac{1}{3}a^3\tau^3\right)\end{aligned} \quad \text{(A4-2)}$$

Substituting for $a,\ b,\ c$ into (A4-2) gives the expression for the approximation to the term structure :

$$R(t,T)=r+\frac{\mu_r}{2}\tau+\frac{1}{6}\left[\sum_{r,x,p}\mu_r\frac{\partial\mu_r}{\partial i}-\sigma_r^2+\frac{1}{2}\sum_{i,j=r,x,p}\rho_{ij}\sigma_i\sigma_j\frac{\partial\mu_r}{\partial i\partial j}\right]\tau^2 \quad \text{(A4-3)}$$

The slope and curvature of the term structure at any point τ along the term structure may be found by taking derivatives :

$$\frac{\partial R(t,T)}{\partial \tau} = \frac{\mu_r}{2} + \frac{1}{3}\left[\sum_{r,x,p} \mu_r \frac{\partial \mu_r}{\partial i} - \sigma_r^2 + \frac{1}{2}\sum_{i,j=r,x,p} \rho_{ij}\sigma_i\sigma_j \frac{\partial \mu_r}{\partial i \partial j}\right]\tau$$
$$\frac{\partial^2 R(t,T)}{\partial \tau^2} = \frac{1}{3}\left[\sum_{r,x,p} \mu_r \frac{\partial \mu_r}{\partial i} - \sigma_r^2 + \frac{1}{2}\sum_{i,j=r,x,p} \rho_{ij}\sigma_i\sigma_j \frac{\partial \mu_r}{\partial i \partial j}\right] \quad \text{(A4-4)}$$

APPENDIX 4-2: FORMULA FOR LAGRANGIAN POLYNOMIAL

A number of methods exist for curve fitting to the term structure. Popular methods include Bernstien polynomials (Schaefer (1981)) or cubic splines (Mastronikola (1991)). Here, interest is for a fit to only to a small number of rates at the short end of the term structure as it is wished to make estimates of the slope and curvature comparable with the approximations in Appendix 4-1. For the purpose at hand, where the number of points for the fit is small, the Lagrange polynomial is quite appropriate.

The Lagrangian polynomial formula allows the fitting of a polynomial of degree n to $n+1$ points on the term structure. From this, information may be extrapolated about non-existent rates, and the slope and curvature of the term structure in the region of the polynomial fitted. The value of the polynomial fit to the term structure at time to maturity τ, is given by :

$$p(t+\tau) = \sum_{i=1\ldots n+1}\left[R(t+\tau_i)\prod_{j=1\ldots n+1;j\neq i}\frac{(\tau-\tau_j)}{(\tau_i-\tau_j)}\right] \quad \text{(A4-5)}$$

The slope and curvature are then given by the first and second derivatives of the expression (A4-5).

$$\frac{\partial p(t+\tau)}{\partial \tau} = \sum_{i=1,2,3}\left[R(t+\tau_i)\prod_{j=1\ldots n+1;j\neq i}\frac{1}{(\tau_i-\tau_j)}\sum_{j=1,2,3;j\neq i}(\tau-\tau_j)\right]$$
$$\frac{\partial^2 p(t+\tau)}{\partial \tau^2} = 2\sum_{i=1,2,3}\left[R(t+\tau_i)\prod_{j=1\ldots n+1;j\neq i}\frac{1}{(\tau_i-\tau_j)}\right] \quad \text{(A4-6)}$$

5. A BETTER APPROACH: THE KALMAN FILTER

5.1 Introduction

The objective of this chapter is to investigate the use the Kalman filter to obtain both the hidden state variables and parameter estimates for interest rate models. In particular, the generalised Vasicek model of Babbs and Nowman (1997) provides a useful basis for the analysis. The two-factor and three factor models developed in chapters 2 and 3 of this thesis are estimated using the techniques developed and explored here. The Kalman filter provides the distinct advantage over the approach taken in chapter 4, in that it will provide both the hidden state variables and the parameter estimates in one compact estimation process.

The chapter proceeds as follows; section 5.2 discusses the development of the Kalman filter and its application to term structure models in finance. Section 5.3 details the general state space form and the algorithm that the Kalman filter entails. The method whereby parameter estimates are found via maximum likelihood is given. A revised method is presented, from the class of Kalman filter derivatives known as '*square-root*' filters. A new method is proposed, based on the '*square-root*' filter, to provide consistent estimates of the analytical derivatives of the log-likelihood function. Section 5.4 applies the Kalman filter to the model of Babbs and Nowman (1997), investigating the case where the exact bond-pricing equation is not available. An approximation to the bond pricing equation is used to approximate the term structure, its slope and curvature at the short end of the yield curve. To investigate the use of this, simulated and empirical analyses are given. It is of particular interest how these approximations perform against their theoretical benchmark as most non-linear models necessitate some form of approximation to the bond pricing equation. Section 5.5 applies the Kalman filter to the two and three factor models developed in chapters 2 and 3. It is shown that the conventional Kalman filter is likely to fail when applied to the potentially chaotic three factor model, whereas the '*square-root*' method consistently provides superior results. Section 5.5 concludes and discusses elements of further work.

5.2 The use of Kalman Filtering: A review

Kalman filtering is applied to a wide range of problems for the estimation of dynamical systems. It is an estimator for the linear quadratic-Gaussian problem; the problem of estimating the state of a linear dynamic system perturbed by Gaussian white noise. Often, it is not possible to measure all the variables of the system that is being investigated. The Kalman filter provides a tool for uncovering these hidden state variables using measurements linearly related to the state. By using estimates linearly related to the state but corrupted by Gaussian white noise, an estimator can be obtained which is optimal in the minimum mean square error sense. For many problems, it is also straightforward to find estimates of the parameters of the model via log-likelihood maximisation.

Many applications have been in control systems with one of the first implementations being the efficient navigation of an Apollo moon mission. These early experiments with the Kalman filter developed such that many navigational estimation and control problems are now handled with the Kalman filter (Schmidt, 1981). Such applications can be seen in Schmidt (1966) and Bucy and Joseph (1968).

More recently, the appeal of Kalman filtering within finance has been recognised. The Kalman filter has proved a popular tool for analysis of term structure models. The nature of the interest rate world means that it is possible to observe a whole term structure of interest rates at any one time. The Kalman filter lends itself well to the application of modelling the time evolution of term structure offering a significant advantage over those techniques modelling only the dynamics of the process or fitting to the term structure on any one day.

In its standard form, it can be applied directly to estimating a number of popular models. This standard form of the Kalman filter requires that the dynamics of the state variables be linear and that the innovations be Gaussian. Term structure models in this class include Vasicek (1977), Langetieg (1980), Babbs and Nowman (1997) and the double decay model of Beaglehole and Tenney (1991). Here, if it is assumed that observed zero coupon yields are measured with error, the model can be put into state space form and the filter applied. Gaussian models have the added advantage of yielding the exact likelihood function from the Kalman filter recursions. This opens the way to consistent and efficient parameter estimates for the model. An initial application of this type was made by Pennachi (1991) and more recently Babbs and Nowman (1997).

Gaussian models represent a significant strand of the literature and are appealing to work with as Kennedy (1994) describes. However, some undesirable properties and restrictions of simple Gaussian models, such as the possibility of negative interest rates, has led to research focusing on other tractable forms for the process. Primarily, this has meant the extension to a class of general exponential-affine models as set out in Duffie and Kan (1996). The application of the Kalman filter has been extended to encompass the case where the process for the state variables is exponential-affine (as in the models of Cox et al (1985b) and Brennan and Schwarz (1980)). Applications of this type include Lund (1997), Jegadeesh and Pennachi (1996) and Ball and Torous (1996).

5.3 Implementation Issues

This section describes the mechanics of the Kalman filter and how the state and parameter estimates may be obtained. A discussion of the '*square-root*' methods is given, describing when they will be preferable to the conventional Kalman filter. A new method is proposed for providing consistent estimates of the analytical derivatives of the log-likelihood function based on the '*square-root*' filter.

5.3.1 The traditional Kalman Filter

Typically, the linear Gaussian model is put into state space form. Following the notation in Harvey(1989) one can write the measurement equation of the state space model.

$$y_t = Z_t(\Psi)\alpha_t + d(\Psi)_t + \varepsilon_t \qquad t = 1,\dots,T \tag{5.3-1}$$

where y_t is an $N\times1$ vector of observables and α_t is an $m\times1$ vector, known as the state vector. When the problem of interest is one of signal extraction then the vector α_t will contain the unobservables. Z_t is an $N\times m$ matrix and d_t is an $N\times1$ vector dependent upon Ψ, the set of parameters for the model. The set of disturbances are serially uncorrelated following[1]

$$\varepsilon_t \sim N(0, H(\Psi)) \tag{5.3-2}$$

The transition equation, determining the time evolution of the vector α_t, is written

$$\alpha_t = T(\Psi)_t\alpha_{t-1} + c(\Psi)_t + \eta_t \qquad t = 1,\dots,T \tag{5.3-3}$$

where the disturbances are distributed in a similar manner as those for the measurement equation. It is assumed that the transition and measurement noise are uncorrelated although this assumption may be relaxed if required. See for instance Park and Kailath (1995) or Harvey (1989).

$$\eta_t \sim N(0, Q(\Psi)) \tag{5.3-4}$$

The Kalman filter recursively calculates the distribution of α_t conditional on the information set at time t for all t up until T. These conditional distributions are normal and as such completely specified by their means and covariances. The Kalman filter propagates these means and covariance matrices. Let a_{t-1} denote the optimal estimator

[1] It is not strictly necessary to assume Gaussian disturbances, however, in the applications made here this assumption will be required.

of α_{t-1} given $\{y_0 \ldots y_{t-1}\}$ Then the optimal estimator of α_t given $\{y_0 \ldots y_{t-1}\}$ is defined to be :

$$a_{t|t-1} = T_t a_{t-1} + c_t \tag{5.3-5}$$

Similarly one can write the expression for the error covariance of the estimate

$$\begin{aligned} P_{t|t-1} &= E\left[\alpha_t - a_{t|t-1}\right]\left[\alpha_t - a_{t|t-1}\right]' \\ &= T_t P_{t-1} T_t' + Q_t \end{aligned} \tag{5.3-6}$$

where P_{t-1} is the $m \times m$ covariance matrix of estimation error.

Together these expressions are known as the predicted estimates. When the new observation of y_t becomes available, the updated (filtered) expressions can be found

$$a_t = a_{t|t-1} + P_{t|t-1} Z_t' F_t^{-1} v_t$$

and

$$P_t = P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1} \tag{5.3-7}$$

where F_t and v_t are

$$\begin{aligned} F_t &= Z_t P_{t|t-1} Z_t' + H_t \\ v_t &= y_t - Z_t a_{t|t-1} - d_t \end{aligned} \tag{5.3-8}$$

The vector v_t is often referred to as the vector of innovations, and F_t is its associated covariance matrix. Often, it is more prudent to invoke one set of recursions moving directly from one predicted estimate to the next. In this format one can write

$$\begin{aligned} a_{t+1|t} &= T_{t+1} a_{t|t-1} + K_t v_t + c_{t+1} \\ P_{t+1|t} &= T_{t+1}\left(P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1}\right) T_{t+1}' + Q_{t+1} \end{aligned} \tag{5.3-9}$$

and the associated matrix, K_t, for the Kalman gain is defined as

$$K_t = T_{t+1} P_{t|t-1} Z_t' F_t^{-1} \tag{5.3-10}$$

Here, one may intuitively think of the gain matrix as representing a weighting factor on the innovations in (5.3-9).

It is these recursions (5.3-9) that will be used and discussed henceforth. For a full discussion of the derivation of the Kalman filter and optimality of the estimators see Harvey(1989) or Grewal and Andrews (1993). The conditional covariance is propagated by its non-linear difference equation in (5.3-9). This is also known as the *Riccati equation* and will be seen to be central to the efficient working of the Kalman filter. In fact the behaviour of the Riccati equation solutions will provide a means for determining the proper running of the filter.

There are two problems that are of primary interest in the application of the Kalman filter. The first is a problem of signal extraction, the second one of parameter estimation. Signal extraction can be performed where knowledge of the dynamical system driving the state variables is available and it is possible to observe measurements linearly related to the system. Whilst the system is linear and the measurement and transition noise are Gaussian, this fits directly to the standard Kalman filter as described above. In the second problem of parameter estimation it is common to use the prediction error decomposition from which it is possible to derive the exact log-likelihood function in the Gaussian case (see Harvey (1989)). The log-likelihood for any given parameter set Ψ is then

$$\log L = -\frac{NT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log|F_t| - \frac{1}{2}\sum_{t=1}^{T} v_t' F_t^{-1} v_t \qquad (5.3\text{-}11)$$

Having computed the likelihood function it can then be maximised with respect to the parameter set Ψ.

5.3.2 Square Root Methods and the Morf-Kailath Filter

Background

Numerical methods will work correctly if implemented with infinite precision arithmetic. Often their theoretical performance is not comparable with their actual performance as evaluated using the precision available on average machines. Early implementations of the Kalman filter found that numerical instabilities were significant. Typically, as a manifestation, the observed mean-squared estimation errors were often much larger than predicted even using simulated data. In some cases, it is possible to find negative values for the error covariance matrix (Grewal and Andrews 1993).

Roundoff errors are symptomatic of the way that computers translate real numbers into floating point arithmetic or fixed point arithmetic. It is prudent to choose both numerical methods and problem formulations that are *robust* against roundoff error. Here, robustness refers to the property that they are not unduly sensitive to error at machine accuracy[2]. An ill-conditioned problem will not be numerically stable with even the most robust implementation methods. For a discussion of the way that roundoff error is created and the effects that it can have see Press et al (1989) and Grewal and Andrews(1993).

A Kalman filtering problem can be considered to be ill-conditioned when the solution of the Riccati equation diverges from the theoretical error covariance matrix of the estimates. If this occurs then the Kalman gain will become sub-optimal. There are a number of reasons for ill-conditioning of the Kalman filtering problem. Poor scaling of the units in the system matrices and poor machine precision are two of the problems that are relatively easy to rectify. Sufficiently accurate measurements and parameter estimates will also affect the performance. Where large errors exist in the values of the system matrices, they are not accounted for in the design of filter. Formulation problems may be another cause of error. Large matrix dimensions imply more arithmetic operations introducing the possibility of more error. Others, such as ill-conditioning in the inversion of F_t may also be difficult to avoid under the standard Kalman filter recursions. One problem, which will be shown to be of particular importance, is the stability of the state transition matrix.

[2] Machine accuracy as defined by Press et al (1989) is "The smallest floating point number which, when added to the floating point number 1.0 produces a number different to 1.0"

In response to the problem of the propagation of numerical errors through the poor solution of the Riccati equation, more numerically stable methods have been proposed that propagate factors of the error covariance matrix. For a survey of different numerical methods and their relative performance see Verhaegen and Van Dooren (1986) or Bierman and Thornton (1977). Both studies find the traditional Kalman filter numerically unreliable with the factorisation algorithms performing significantly better. In particular the Kalman filter is particularly susceptible to the condition number of H_t and the eigenvalues of T_t as Verhaegen and Van Dooren (1986) show.

The condition number of a matrix A may be defined as (Golub and Van Loan, 1983)[3]

$$\kappa(A) = \|A\|_2 \|A^{-1}\|_2 = \sigma_1(A) / \sigma_n(A) \tag{5.3-12}$$

where σ_1 and σ_n are the largest and smallest singular values respectively. The condition number will always be greater than 1. In general, matrices with small condition numbers (close to 1) will be considered well conditioned whilst large condition numbers are a cause for concern and the matrix will be considered ill-conditioned. In practice, this concept refers to numerical stability of matrix operations involving the matrix A. In respect of the Kalman filter the condition number of H_t is relevant in the construction of F_t which is inverted in the matrix Riccati equation. Park and Kailath (1995) suggest that it is beneficial, not only to ensure the positive definiteness of H_t, but also to attempt to move it as close to the identity matrix as possible.

For the transition equation and hence the Kalman filter to be stable it is required that all the eigenvalues of the matrix T_t are less than one in absolute value (see, for example Hamilton (1989)). That is

[3] The sensitivity of the solution x to the linear system problem $Ax=b$ is quantified by the *condition number* of the matrix A. If A is well conditioned then the condition number is small, if it is singular then the condition number is ∞. For a discussion concerning different methods of calculating the condition number see Golub and Van Loan (1983).

$$|\lambda_i(T_t)| < 1, \qquad i = 1, \ldots, m \tag{5.3-13}$$

This is a general property of systems of difference equations (see for instance Devaney (1989) or Simon and Blume (1994)). Verhaegen and Van Dooren (1986) show that the conventional Kalman filter can fail in the case where the largest eigenvalue of T_t has absolute value 1. In the majority of cases of interest in economics and finance, it is a desirable property of the model that the transition matrix is stable. The potentially chaotic three factor model developed in chapter 3 is not stable in the traditional sense and it will be seen that this will pose a problem for the application of the traditional Kalman filter.

A number of different methods have been developed which propagate '*square-root*' factors[4] of the error covariance matrix (or its inverse). The advantage comes directly in that the square will always be non-negative definite. It will also be better conditioned as the condition number for the matrix $P_{t|t-1}$ will always be the square of that of $P_{t|t-1}^{1/2}$. These methods are however often more computationally intensive, although in some cases, such as when the system matrices are time-invariant, they can be made as or more efficient than the original Kalman filter (Verhaegen and Van Dooren (1986)).

The Morf-Kailath/Square Root Covariance Filter

This algorithm was first introduced by Morf and Kailath (1975) providing a compact and efficient method for a combined temporal and observational update of the Cholesky factor of the error covariance matrix. Further refinements can be found in Park and Kailath (1995) although the basic method remains unaltered. The '*square-root*' covariance filter (SRCF) addresses the numerical instabilities that are discussed for the conventional Kalman filter (conventional KF) above. The SRCF algorithm is now presented.

[4] These factors are not true square roots but in fact Cholesky factors. However, the term '*square root*' has pervaded and will be used interchangeably with that of Cholesky factor.

The Cholesky factor C_A of a matrix A is defined as having the property $C_A C_A' = A$. Any symmetric non-negative definite matrix will have Cholesky factors, which will not be unique. With suitable constraints on the form of the factor a unique C_A can be found. Typically, it is required that the Cholesky factors are either upper triangular or lower triangular in form[5].

The algorithm proceeds recursively as follows. At time t the Cholesky factors of the measurement and transition noise, C_{H_t} and C_{Q_t} respectively, are found[6]. The Cholesky factor of the error covariance need only be found for the first recursion, whence it will be propagated recursively. Form the *pre-array* A_t

$$A_t = \begin{bmatrix} C_{Q_t} & T_{t+1}C_{P_{t|t-1}} & 0 \\ 0 & Z_t C_{P_{t|t-1}} & C_{H_t} \end{bmatrix} \tag{5.3-14}$$

The symmetric product A_tA_t' has the form

$$A_tA_t' = \begin{bmatrix} T_{t+1}P_{t|t-1}T_{t+1}' + Q_{t+1} & T_{t+1}P_{t|t-1}Z_t' \\ Z_tP_{t|t-1}T_{t+1}' & Z_tP_{t|t-1}Z_t' + H_t \end{bmatrix} \tag{5.3-15}$$

The QR decomposition enables any general matrix to be decomposed into a product of an orthogonal and triangular matrix. Thus one may write[7]

$$A_t = C_{t+1|t}\Theta_t \tag{5.3-16}$$

where $C_{t+1|t}$ is triangular and Θ_t is orthogonal. Hence :

$$\begin{aligned} A_tA_t' &= \left(C_{t+1|t}\Theta_t\right)\left(C_{t+1|t}\Theta_t\right)' \\ &= C_{t+1|t}\Theta_t\Theta_t'C_{t+1|t}' \\ &= C_{t+1|t}\left(\Theta_t\Theta_t'\right)C_{t+1|t}' \\ &= C_{t+1|t}C_{t+1|t}' \end{aligned} \tag{5.3-17}$$

[5] A matrix is called upper triangular if its elements on and above the diagonal are non-zero, all elements below are zero. The reverse is true for a lower triangular matrix.

[6] See Nash (1979) or Grewal and Andrews (1993) for Cholesky factorisation algorithms.

[7] Here we are really observing the transposed form of $A' = \Theta_t' C_{P_{t+1|t}}'$

so that $C_{t+1|t}$ is a triangular Cholesky decomposition of $A_t A_t'$. Following from this it can be seen that if our pre-array A_t is upper triangularised in the form $A_t \Theta_t' = C_{t+1|t}$ then $C_{t+1|t}$ is obliged to be of the form :

$$C_{t+1|t} = \begin{bmatrix} 0 & C_{P_{t+1|t}} & \psi_t \\ 0 & 0 & C_{F_t} \end{bmatrix} \tag{5.3-18}$$

where the elements of the *post-array,* $C_{t+1|t}$, conform to

$$\begin{aligned} C_{F_t} C_{F_t}' &= Z_t P_{t|t-1} Z_t' + H_t \\ &= F_t \\ \psi_t &= T_{t+1} P_{t|t-1} Z_t' C_{F_t}^{-1'} \\ C_{P_{t+1|t}} C_{P_{t+1|t}}' &= T_{t+1} P_{t|t-1} T_{t+1}' + Q_{t+1} - \psi_t \psi_t' \\ &= P_{t+1|t} \end{aligned} \tag{5.3-19}$$

The elements $C_{P_{t+1|t}}$ and C_{F_t} are the Cholesky factors of $P_{t+1|t}$ and F_t respectively. The Kalman gain can be computed from the elements of the post array

$$K_t = \psi_t C_{F_t}^{-1} \tag{5.3-20}$$

In practice the inverse of the Cholesky factor of F_t need not be computed and the Kalman gain can be obtained by solving equation (5.3-20) by back substitution. The error-covariance matrix $P_{t+1|t}$ need never be computed (if not desired). The triangularisation method employed may be a standard method such as Givens rotations or Householder reflections. See for instance Press et al (1989), Nash (1979), Grewal and Andrews (1993) and Golub and Van Loan (1983). An implementation based upon that given in Golub and Van Loan (1983) (see this text for details) is given in the appendix. A number of other methods exist which may be numerically superior. See Park and Kailath (1995) for a discussion of these.

The SRCF has several appealing properties over the conventional KF recursions which can be summarised below :

- The SRCF propagates '*square-root*' factors of the error covariance matrix; these factors will be better conditioned numerically and ensure that the error covariance matrix itself is at least positive semi-definite.
- The error covariance matrix need never be calculated if not desired.
- Propagation is achieved via the '*square-root*' factors.
- The temporal/observational update is effected by an orthogonal transformation. The roundoff properties of this method are very favourable (see Golub and Van Loan (1983))
- The inversion of F_t comes for free.
- Numerical studies show '*square-root*' factorisation methods to be stable in a variety of circumstances where the conventional KF will fail. This includes cases where the spectral radius of T_{t+1} is greater than or equal to 1.

5.3.3 Consistent propagation of derivatives

Estimates of the derivatives of the log-likelihood function with respect to the elements of the parameter set Ψ are required for two uses; standard error computation and efficient log-likelihood maximisation routines. Firstly, following Babbs and Nowman (1997), when computing maximum-likelihood estimates of the elements of the parameter set Ψ it is desired to compute the associated standard errors from the formula in Hamilton (1994). This is done by computing the *information matrix* $I_{ij}(\Psi)$, formed from the second derivatives of the likelihood function. This converges to a positive definite matrix when scaled by (1/T). The reported standard errors for Ψ_T are then calculated as the square roots of the diagonal elements of $(1/T)[I_{ij}(\Psi_T)/T]^{-1}$.

Secondly, the derivatives of the log-likelihood function may be wanted for use in function maximisation routines which require the derivatives. In particular, the popular BFGS variant of the Davidson Fletcher Powell routine is such a case. See for example

Press et al (1989) or Hamilton (1989). The particular forms of the information matrix (5.3-22) and score vector (5.3-21) of the log-likelihood function can be seen to be (Harvey (1989, p142))

$$\frac{\partial \log L}{\partial \Psi_i} = -\frac{1}{2}\sum_t \left\{ \mathrm{tr}\left[\left[F_t^{-1}\frac{\partial F_t}{\partial \Psi_i}\right]\left(I - F_t^{-1}v_t v_t'\right)\right] + 2\frac{\partial v_t'}{\partial \Psi_i}F_t^{-1}v_t \right\}, \quad i = 1,\ldots,n \tag{5.3-21}$$

$$\begin{aligned} I_{ij}(\Psi) &= \frac{1}{2}\sum_t \left[\mathrm{tr}\left[F_t^{-1}\frac{\partial F_t}{\partial \Psi_i}F_t^{-1}\frac{\partial F_t}{\partial \Psi_j}\right]\right] \\ &\quad + E\left[\sum_t \left(\frac{\partial v_t'}{\partial \Psi_i}\right)' F_t^{-1}\frac{\partial v_t}{\partial \Psi_j}\right], \quad i,j = 1,\ldots,n \end{aligned} \tag{5.3-22}$$

In computing the information matrix it is practical to assume that the limit of $I_{ij}(\Psi)$ as $T \to \infty$ is the same as the plim of

$$I_{ij}(\Psi) = \frac{1}{2}\sum_t \left[\mathrm{tr}\left[F_t^{-1}\frac{\partial F_t}{\partial \Psi_i}F_t^{-1}\frac{\partial F_t}{\partial \Psi_j}\right] + \left(\frac{\partial v_t'}{\partial \Psi_i}\right)' F_t^{-1}\frac{\partial v_t}{\partial \Psi_j}\right]\Bigg|_{\Psi=\hat{\Psi}_T} \qquad i,j = 1,\ldots,n \tag{5.3-23}$$

Computation of the derivatives may be done either numerically or analytically.

Numerical computation of the derivatives has several drawbacks. It is a feature of numerically computed derivatives that they are highly susceptible to machine accuracy. See for example Nash (1979) for a discussion of this. Furthermore, the more convoluted the function in question, the more the scope for error in the computation of the derivatives. Highly non-linear functions and cases of discontinuities can be detrimental to the process of calculating numerical derivatives.

If only the score vector is required then the derivatives of the log-likelihood function may be evaluated directly. Where the information matrix is desired then the derivatives $\partial F_t / \partial \Psi_i$ and $\partial v_t / \partial \Psi_i$ are computed. In the case of F_t, it is required to evaluate

$$\partial F_t(\Psi)/\partial \Psi_i = \lim_{h \to 0}\left[F_t(\Psi + he_i) - F_t(\Psi)\right]/h \qquad (5.3\text{-}24)$$

where e_i is the ith column of a unit matrix of order np (where np is the number of parameters in Ψ). As many such derivative evaluations need to be made in the course of maximising the function, methods for evaluating (5.3-24) by means of function interpolation formulae are too slow to be used in this case. In practice, (5.3-24) may be approximated by the expression

$$\partial F_t(\Psi)/\partial \Psi_i = \left[F_t(\Psi + he_i) - F_t(\Psi)\right]/h \qquad (5.3\text{-}25)$$

where the problem becomes one of choosing a value for h. This is not a trivial problem; some method of finding a value for h which is numerically accurate is required. If h is chosen too large, then the line implied by the slope will not be tangential to the function. If h is chosen too small then the accuracy with which the computer stores the values may imply a degree of digit cancellation leading to inaccuracy in the calculation of the derivative.

Nash (1979) suggests a method for determining the value of h so as to minimise such errors. It is suggested setting

$$h = \left(|\Psi_i| + \varepsilon^{1/2}\right)\varepsilon^{1/2} \qquad (5.3\text{-}26)$$

where ε is machine precision. h will be different for each parameter axis. The value for h cannot become smaller than machine precision. As it is scaled by the value of the parameter, it will always change at least the right most half of the digits in Ψ_i.

Computation of the derivatives analytically involves running a parallel set of recursions for the derivatives of the estimate $a_{t|t-1}$ and the covariance matrix of estimation error $P_{t|t-1}$. From this one can derive the required values necessary to form the score vector and information matrix (5.3-21) and (5.3-22) respectfully. Harvey (1989, p143) provides these recursions based upon the conventional KF recursions.

Following from the previous discussion surrounding the numerical stability properties of the conventional KF recursions one can expect the recursions of the derivatives to be similarly conditioned. In forming the score vector and information matrix, the derivatives of F_t and v_t with respect to Ψ are required.

$$\begin{aligned}
\frac{\partial v_t}{\partial \Psi_i} &= -Z_t \frac{\partial a_{t|t-1}}{\partial \Psi_i} - \frac{\partial Z_t}{\partial \Psi_i} a_{t|t-1} - \frac{\partial d_t}{\partial \Psi_i} \\
\frac{\partial F_t}{\partial \Psi_i} &= \frac{\partial Z_t}{\partial \Psi_i} P_{t|t-1} Z_t' + Z_t \frac{\partial P_{t|t-1}}{\partial \Psi_i} Z_t' + Z_t P_{t|t-1} \frac{\partial Z_t'}{\partial \Psi_i} + \frac{\partial H_t}{\partial \Psi_i}
\end{aligned} \tag{5.3-27}$$

These quantities are reliant upon the recursions for the derivatives of the predicted estimates of $a_{t+1|t}$ and $P_{t+1|t}$.

$$\begin{aligned}
\frac{\partial a_{t+1|t}}{\partial \Psi_i} &= \frac{\partial T_{t+1}}{\partial \Psi_i} a_{t|t-1} + T_{t+1} \frac{\partial a_{t|t-1}}{\partial \Psi_i} + \frac{\partial K_t}{\partial \Psi_i} v_t + K_t \frac{\partial v_t}{\partial \Psi_i} + \frac{\partial c_{t+1}}{\partial \Psi_i} \\
\frac{\partial P_{t+1|t}}{\partial \Psi_i} &= \frac{\partial T_{t+1}}{\partial \Psi_i} P_t T_{t+1}' + T_{t+1} \frac{\partial P_t}{\partial \Psi_i} T_{t+1}' + T_{t+1} P_t \frac{\partial T_{t+1}'}{\partial \Psi_i} + \frac{\partial Q_{t+1}}{\partial \Psi_i}
\end{aligned} \tag{5.3-28}$$

where the derivatives of P_t and K_t are calculated as follows

$$\begin{aligned}
\frac{\partial P_t}{\partial \Psi_i} &= \frac{\partial P_{t|t-1}}{\partial \Psi_i} - \frac{\partial P_{t|t-1}}{\partial \Psi_i} Z_t' F_t^{-1} Z_t P_{t|t-1} - P_{t|t-1} \frac{\partial Z_t'}{\partial \Psi_i} F_t^{-1} Z_t P_{t|t-1} - P_{t|t-1} Z_t' \frac{\partial F_t^{-1}}{\partial \Psi_i} Z_t P_{t|t-1} \\
&\quad - P_{t|t-1} Z' F_t^{-1} \frac{\partial Z_t}{\partial \Psi_i} P_{t|t-1} - P_{t|t-1} Z' F_t^{-1} Z_t \frac{\partial P_{t|t-1}}{\partial \Psi_i} \\
\frac{\partial K_t}{\partial \Psi_i} &= \frac{\partial T_{t+1}}{\partial \Psi_i} P_{t|t-1} Z_t' F_t^{-1} + T_{t+1} \frac{\partial P_{t|t-1}}{\partial \Psi_i} Z_t' F_t^{-1} + T_{t+1} P_{t|t-1} \frac{\partial Z_t'}{\partial \Psi_i} F_t^{-1} + T_{t+1} P_{t|t-1} Z_t' \frac{\partial F_t^{-1}}{\partial \Psi_i}
\end{aligned} \tag{5.3-29}$$

It is a natural progression to think that a factorisation technique could be employed for the recursions of the derivatives in a similar manner to those for the straight values of the Kalman filter. Upon inspection of the method of the Morf-Kailath filter, it can readily be seen that the filter works by exploiting the natural symmetry exhibited by the expressions for F_t and $P_{t+1|t}$. It would be useful if it were possible to calculate the Cholesky factors of the derivatives of these quantities. In this way, it would be possible to proceed calculating the recursions of the Cholesky factors of the derivatives for F_t

and $P_{t+1|t}$. However, whilst it is always the case that the derivative of a covariance matrix (such as $P_{t+1|t}$) will be symmetric, nothing can be said about its definiteness. Although its definiteness is ambiguous in this case, it is still useful to be able to ensure that the covariance matrices derivatives are symmetric. From the standard Kalman filter recursions, it is the case that a loss of symmetry in the covariance matrix of estimation error is symptomatic of numerical instability and measures taken to rectify this will be beneficial to the filter performance.

A method to ensure symmetry in covariance matrix derivatives.

Assuming that the Morf-Kailath filter is used, this opens the way to a method of calculating the recursions for the derivatives of the quantities of the Kalman Filter such that the symmetry of the derivatives of F_t and $P_{t+1|t}$ is assured. As already noted, the Morf-Kailath filter makes use of the natural symmetry in the recursions, which further allows a procedure for exploiting symmetry in the recursions for the derivatives. Here, I propose a procedure detailing how the quantities formed during the application of the Morf-Kailath filter may be used to obtain symmetrical derivatives of the covariance matrix of estimation error and the covariance matrix of the innovations. To my knowledge, this is the first algorithm implemented with the motivation of producing analytical derivatives of the log-likelihood function which are consistent with the numerical properties of the SRCF

From the quantities used to form the pre-array (5.3-14) define :

$$\xi_t = Z_t C_{P_{t|t-1}} \tag{5.3-30}$$

$$\zeta_t = T_{t+1} C_{P_{t|t-1}} \tag{5.3-31}$$

so that ψ_t from the post-array (5.3-18) can be otherwise expressed as

$$\psi_t = T_{t+1} P_{t|t-1} Z_t' C_{F_t}^{-1'} = \zeta_t \xi_t' C_{F_t}^{-1'}$$

Under these definitions, the recursion for $P_{t+1|t}$ (5.3-9) can be re-expressed as :

$$P_{t+1|t} = \zeta_t \zeta_t' - \psi_t \psi_t' + Q_{t+1} \tag{5.3-32}$$

Taking the derivative of this with respect to parameter Ψ_i.

$$\begin{aligned}
\frac{\partial P_{t+1|t}}{\partial \Psi_i} &= \frac{\partial\left(\zeta_t \zeta_t'\right)}{\partial \Psi_i} - \frac{\partial\left(\psi_t \psi_t'\right)}{\partial \Psi_i} + \frac{\partial Q_{t+1}}{\partial \Psi_i} \\
&= \frac{\partial \zeta_t}{\partial \Psi_i}\zeta_t' + \frac{\partial \zeta_t'}{\partial \Psi_i}\zeta_t - \left[\frac{\partial \psi_t}{\partial \Psi_i}\psi_t' + \frac{\partial \psi_t'}{\partial \Psi_i}\psi_t\right] + \frac{\partial Q_{t+1}}{\partial \Psi_i} \\
&= \frac{\partial \zeta_t}{\partial \Psi_i}\zeta_t' + \left(\frac{\partial \zeta_t}{\partial \Psi_i}\zeta_t'\right)' - \left[\frac{\partial \psi_t}{\partial \Psi_i}\psi_t' + \left(\frac{\partial \psi_t}{\partial \Psi_i}\psi_t'\right)'\right] + \frac{\partial Q_{t+1}}{\partial \Psi_i}
\end{aligned} \tag{5.3-33}$$

(5.3-33) is a much more convenient form to manipulate than the equivalent combined expressions in (5.3-28) and (5.3-29) for the recursion of the derivative of $P_{t+1|t}$ from one time period to the next. This is due mainly to the efficient usage of the matrices ξ_t, ζ_t and ψ_t, which are shared between the SRCF and the derivative algorithm. In a similar manner the derivative for F_t can be formed.

$$\begin{aligned}
\frac{\partial F_t}{\partial \Psi_i} &= \frac{\partial\left(\xi_t \xi_t'\right)}{\partial \Psi_i} + \frac{\partial H_t}{\partial \Psi_i} \\
&= \frac{\partial \xi_t}{\partial \Psi_i}\xi_t' + \frac{\partial \xi_t'}{\partial \Psi_i}\xi_t + \frac{\partial H_t}{\partial \Psi_i} \\
&= \frac{\partial \xi_t}{\partial \Psi_i}\xi_t' + \left(\frac{\partial \xi_t}{\partial \Psi_i}\xi_t'\right)' + \frac{\partial H_t}{\partial \Psi_i}
\end{aligned} \tag{5.3-34}$$

It can be readily seen that the quantities formed in (5.3-33) and (5.3-34) are symmetric by construction. In practice it is unlikely that symmetry will be lost in their construction as the operation of adding a matrix to its transpose should be well conditioned. It is assumed that forming the derivatives of the system matrices H_t and Q_{t+1} should be done so as to as to assure they are symmetric.

It is still left to define the derivatives of the quantities ξ_t, ψ_t, and ζ_t if the above expressions are to be implemented.

$$\frac{\partial \zeta_t}{\partial \Psi_i} = \frac{\partial T_{t+1}}{\partial \Psi_i} CP_{t|t-1} + T_{t+1} \frac{\partial CP_{t|t-1}}{\partial \Psi_i}$$

$$\frac{\partial \xi_t}{\partial \Psi_i} = \frac{\partial Z_t}{\partial \Psi_i} CP_{t|t-1} + Z_t \frac{\partial CP_{t|t-1}}{\partial \Psi_i} \tag{5.3-35}$$

$$\frac{\partial \psi_t}{\partial \Psi_i} = \frac{\partial \zeta_t}{\partial \Psi_i} \xi_t CF_t^{-1'} + \zeta_t \frac{\partial \xi_t}{\partial \Psi_i} CF_t^{-1'} + \zeta_t \xi_t \frac{\partial CF_t^{-1'}}{\partial \Psi_i}$$

Here, a complication arises in that it is necessary to construct the derivative of a Cholesky factor of a matrix. I propose here a method whereby these quantities can be uncovered. Providing the derivative of the original matrix (to which the Cholesky factor relates) and the triangular Cholesky factor itself, are available, then the following procedure may be used.

For any symmetric positive definite matrix A, the expression (5.3-36) can be solved to give the derivative of the triangular Cholesky factor of A, $\partial C_A / \partial \Psi_i$.

$$\frac{\partial A}{\partial \Psi_i} = \frac{\partial C_A}{\partial \Psi_i} C_A{}' + C_A \frac{\partial C_A{}'}{\partial \Psi_i} \tag{5.3-36}$$

If C_A and $\partial C_A/\partial \psi_i$ are upper triangular $n \times n$ matrices, $\partial A/\partial \psi_i$ $n \times n$ symmetric, then (5.3-36) will be of the form

$$\begin{bmatrix} \partial A_{11} & \partial A_{12} & \cdots & \partial A_{1n} \\ \partial A_{12} & \partial A_{22} & & \partial A_{2n} \\ \vdots & & \ddots & \vdots \\ \partial A_{1n} & \partial A_{2n} & \cdots & \partial A_{nn} \end{bmatrix} = \begin{bmatrix} \partial C_{A_{11}} & \partial C_{A_{12}} & \cdots & \partial C_{A_{1n}} \\ 0 & \partial C_{A_{22}} & & \partial C_{A_{2n}} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \partial C_{A_{nn}} \end{bmatrix} \times \begin{bmatrix} C_{A_{11}} & 0 & \cdots & 0 \\ C_{A_{12}} & C_{A_{22}} & & \vdots \\ \vdots & & \ddots & 0 \\ C_{A_{1n}} & C_{A_{2n}} & \cdots & C_{A_{nn}} \end{bmatrix}$$
$$+ \left[\begin{bmatrix} \partial C_{A_{11}} & \partial C_{A_{12}} & \cdots & \partial C_{A_{1n}} \\ 0 & \partial C_{A_{22}} & & \partial C_{A_{2n}} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \partial C_{A_{nn}} \end{bmatrix} \times \begin{bmatrix} C_{A_{11}} & 0 & \cdots & 0 \\ C_{A_{12}} & C_{A_{22}} & & \vdots \\ \vdots & & \ddots & 0 \\ C_{A_{1n}} & C_{A_{2n}} & \cdots & C_{A_{nn}} \end{bmatrix} \right]^T$$

For convenience, the notation ∂A, will be used to denote the use/construction of all the matrices $\partial A / \partial \Psi_i$ for $i = 1...n$. A unique value can be found for each of the elements $\partial C_A/\partial \psi_i$. I have been unable to find elsewhere in the literature any reference to this or a similar algorithm which enables the uncovering of the derivative of the Cholesky factor.

Here I propose how this may be effected for the cases where the Cholesky factors are either upper or lower triangular.

ALGORITHM 5-1: UNCOVERING THE DERIVATIVE OF A UPPER TRIANGULAR CHOLESKY FACTOR

Given the derivative of a symmetric matrix, $D \in \mathbf{R}^{n \times n}$, and the Cholesky factor of the original matrix, $C \in \mathbf{R}^{n \times n}$, the derivative of the Cholesky factor, $B \in \mathbf{R}^{n \times n}$ can be uncovered. C, B, upper triangular.

$m := \max\{|C_{ij}|\} \quad i = n...1, j = n...i$

For $i = n$ downto 1

For $j = n$ downto i

$D_{ij} := D_{ij}/m$

$C_{ij} := C_{ij}/m$

If $i = j$ Then $B_{ij} := \left[D_{ij} - \left(\sum_{k=j+1..n} B_{ik} C_{jk} + \sum_{k=j+1..n} C_{ik} B_{jk} \right) \right] \Big/ 2C_{jj}$

Else $B_{ij} := \left[D_{ij} - \left(\sum_{k=j+1..n} B_{ik} C_{jk} + \sum_{k=j..n} C_{ik} B_{jk} \right) \right] \Big/ C_{jj}$

For $i = 1$ to n

For $j = 1$ to i-1

$B_{ij} := 0$

ALGORITHM 5-2: UNCOVERING THE DERIVATIVE OF A LOWER TRIANGULAR CHOLESKY FACTOR

Given the derivative of a symmetric matrix, $D \in \mathbf{R}^{n \times n}$, and the Cholesky factor of the original matrix, $C \in \mathbf{R}^{n \times n}$, the derivative of the Cholesky factor, $B \in \mathbf{R}^{n \times n}$ can be uncovered. C, B, lower triangular.

$m := \max\{|C_{ij}|\} \quad i = 1...n, j = 1...i$

For $i = 1$ to n

For $j = 1$ to i

$D_{ij} := D_{ij}/m$

$C_{ij} := C_{ij}/m$

If $i = j$ Then $B_{ij} := \left[D_{ij} - \left(\sum_{k=1..j-1} B_{ik} C_{jk} + \sum_{k=1..j-1} C_{ik} B_{jk} \right) \right] \Big/ 2C_{jj}$

Else $B_{ij} := \left[D_{ij} - \left(\sum_{k=1..j-1} B_{ik} C_{jk} + \sum_{k=1..j} C_{ik} B_{jk} \right) \right] \Big/ C_{jj}$

For $j = 1$ to n

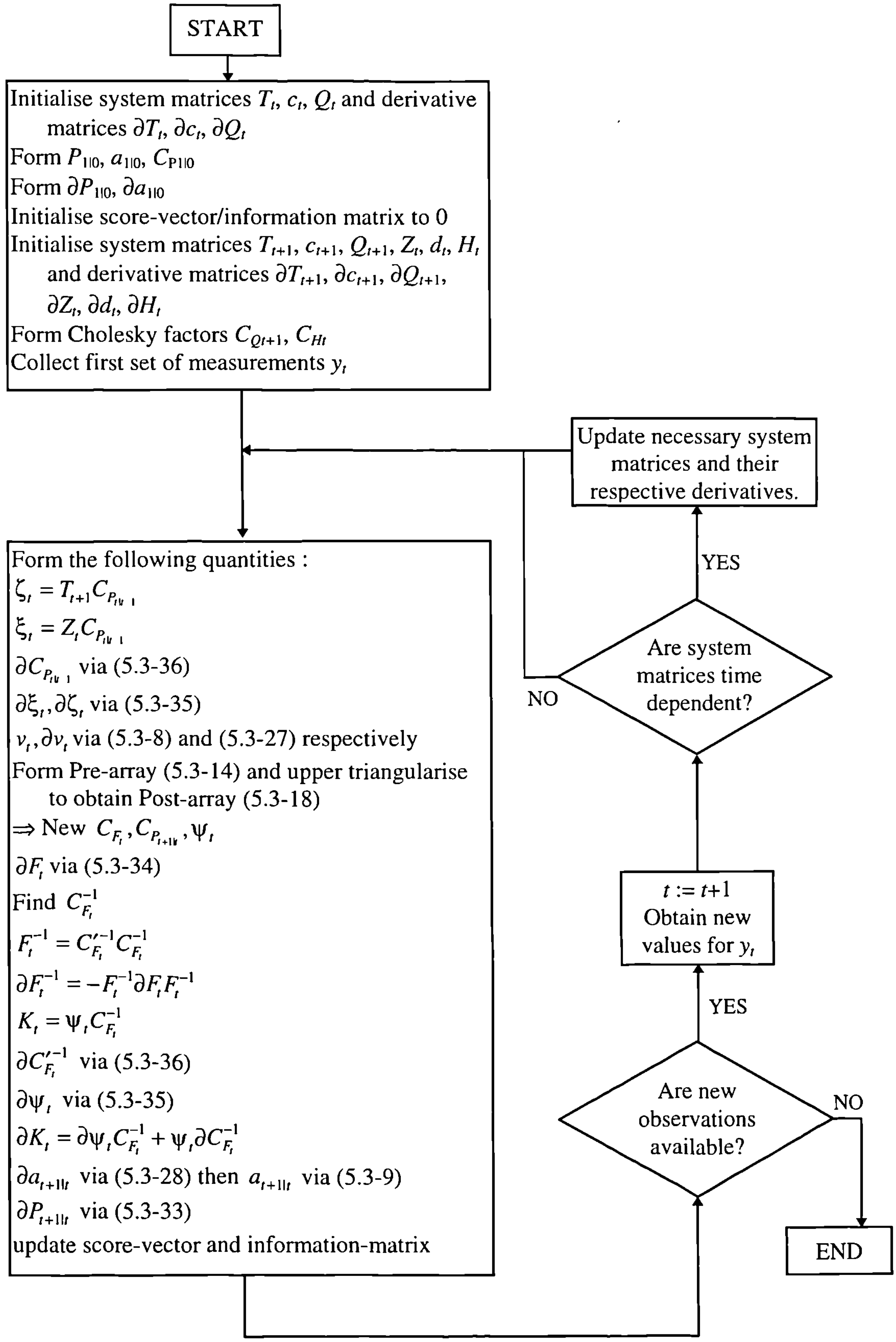
For $i = 1$ to j-1

$B_{ij} := 0$

Pascal coding for algorithms Algorithm 5-1 and Algorithm 5-2 is presented in the appendix to this chapter, for the upper triangular form and lower triangular forms. It is noted that the numerical roundoff properties of these algorithms may be improved by better scaling and/or iterative procedures. The lower triangular form will be required when solving for the derivative of the inverse transpose of the Cholesky factor of F_t. The upper triangular version is required for solving for the derivative of the Cholesky factor of $P_{t|t-1}$.

The full procedure for the recursions required to calculate the analytical derivatives of the Kalman filter recursions, such that the score-vector and information-matrix can be obtained are summarised in the flow diagram Algorithm 5-3 below.

ALGORITHM 5-3: CALCULATION OF SCORE-VECTOR/INFORMATION-MATRIX FOR SRCF USING ANALYTICAL DERIVATIVES



Algorithm 5-3 has several pleasing properties.

- It is based on the SRCF filter, with all the benefits arising from that
- Many of the quantities required are already formed as part of the SRCF filter.
- Only one matrix inversion need be performed. This is the inversion of the Cholesky factor of F_t. As the condition number of the Cholesky factor is the square-root of that of the original matrix, this operation can be considered well conditioned, providing a method for inversion such as LU decomposition and back-substitution is used.
- The derivatives of the covariance matrices $P_{t|t-1}$ and F_t are symmetric by construction and the operations involved in forming them are unlikely to cause loss of symmetry in practice.

Given the widespread use of the '*square-root*' methods in the engineering literature it seems somewhat surprising that little reference is given to it in texts and applications with the finance and economics area. Harvey (1989) notes that

> *"The use of the square-root algorithm in the filter is generally regarded as the most numerically stable algorithm. Despite this attraction, square root filters have not been used a great deal outside engineering, although an exception is Kitagawa (1981)."*

Hamilton (1994) notes that when performing log-likelihood maximisation, the numerical search will be better behaved if the error-covariance matrix of the estimates is parameterised in terms of its Cholesky factorisation. This is most likely to be true where there is the possibility of the maximisation routine choosing parameters for the model which would lead the Kalman filter into an unstable region. However, no treatment is given to '*square-root*' filters.

It may be true that most finance/economics applications are constrained at the modelling stage to have stable properties. If parameter estimation is to be performed using the conventional KF, care should be taken not to let the maximisation method choose parameter estimates in an unstable region as noted above. Verhagen and Van Dooren (1986) find that

> *"when the accuracy of the Kalman gain is considered no preference should exist for the square-root filters to the conventional KF when* T_t *is* ***stable and time invariant****. (For situations where* T_t *has eigenvalues* ***on or outside*** *the unit circle, the conventional KF has to be changed, e.g. to the conventional KF(*S*) implementation). However, the experimental results demonstrate that for the latter conditions the loss of accuracy with a conventional KF(S) is still higher than the square-root filters."*

The conventional KF(*S*) variant to which Verhagen and Van Dooren refer incorporates an attempt to rectify loss of symmetry in conventional KF error-covariance matrices. This may be done for instance, by averaging the off-diagonal elements, or replicating the upper triangular part of the matrix with its lower triangular half.

From the discussion presented in this section, I would suggest that it is prudent to, wherever possible, use a '*square-root*' variant of the Kalman filter. Algorithms such as the SRCF presented above are compact and require little extra programming for the benefits they entail.

5.4 The Babbs and Nowman (1997) model and use of approximated term structure.

5.4.1 Introduction

This section seeks to explore the application of Kalman filtering to the Babbs and Nowman (1997) model. In their paper they develop a subclass of the generalised

Vasicek model after Langetieg (1980). They apply the Kalman filter to one and two factor models using US data on eight maturites. The exact discretisation of the process for the state variables and the bond-pricing equation are found allowing them to put the model into the state-space form for the transition and measurement equation. Here, an investigation is undertaken into the stability of this under the conventional KF and also the case where the exact bond-pricing equation is not known, but approximated.

5.4.2 The Babbs and Nowman (1997) model

Babbs and Nowman (1997) develop a subclass of the general linear Gaussian after Langetieg (1980). The model assumes no dependence of the volatility of rates on the level. Further, their subclass is characterised by the drift of each state variable depending only upon the state variable to which it relates. This is particularly beneficial in that an explicit formula can be found relating pure discount bond prices to the state-variables. These state variables are not directly allied with observable quantities such as spot rates.

The General Linear Gaussian Model and the B&N subclass.

The expressions (5.4-1) and (5.4-2) represent the dynamics of the state variables and the general linear combination of the state variables giving the short rate respectively.

$$dX_j = \left(a_j + \sum_{k=1}^{J} B_{jk} X_k \right) dt + c_j dW_j \qquad (5.4\text{-}1)$$

$$r = w_0 + \sum_{j=1}^{J} w_j X_j \qquad (5.4\text{-}2)$$

For the Babbs and Nowman subclass, (5.4-1) and (5.4-2) are restricted to be of the following form where μ is the long term mean of the process for r:

$$dX_j = -\xi_j X_j dt + c_j dW_j \qquad (5.4\text{-}3)$$

$$r = \mu - \sum_{j=1}^{J} X_j \qquad (5.4\text{-}4)$$

In each of the two cases above $X_1,\ldots,X_J$ are the unobservable state variables of the model. Babbs and Nowman suggest that in a two factor model, the first factor may represent the effect of short-term economic "news" such as market rumours and interest rate decisions from the monetary authority. The second factor would then relate to long-term economic news as provided by monthly/quarterly economic reports. The c_j are deterministic processes acting as the diffusion coefficients. The $W_1,\ldots,W_J$ are standard Brownian motions with instantaneous correlation processes ρ_{jk} for $j,k=1,\ldots,J$.

As Babbs and Nowman (1997) detail, their subclass involves three essential restrictions on the form of (5.4-1) and (5.4-2)

- The matrix of mean reversion coefficients, B_{jk}, has the off-diagonal terms restricted to be zero.
- The level coefficient functions, a_j, have been set identically to zero.
- The weighting coefficient functions on the state variables in the expression for the instantaneous short rate are set to minus unity.

Their first restriction is motivated by analytical tractability in that a closed form solution to the bond pricing formula can be found. They then show that the other two restrictions are not, as such, restrictions, but follow naturally, allowing for the removal of unnecessary parameters from the model.

For the case where the parameters of the model are constant, Babbs and Nowman present the formula for the price of a pure discount bond. This evaluates to :

$$B(M,t)=\exp\left\{-\tau\left[R(\infty)-w(\tau)-\sum_{j=1}^{J}H(\xi_j\tau)X_j(t)\right]\right\} \tag{5.4-5}$$

where

$$\tau \equiv M-t$$
$$H(x)=\frac{1-e^{-x}}{x} \tag{5.4-6}$$

$B(M,t)$ represents the price at time t of a unit nominal pure discount bond maturing at time M. Here, the cases of interest are for J=1 and J=2, where the quantities $R(\infty)$ and $w(\tau)$ evaluate to

For $J = 1$

$$R(\infty) = \mu + \frac{\theta_1 c_1}{\xi_1} - \frac{1}{2}\left(\frac{c_1}{\xi_1}\right)^2 \tag{5.4-7}$$

$$w(\tau) = H(\xi_1\tau)\left[\frac{\theta_1 c_1}{\xi_1} - \left(\frac{c_1}{\xi_1}\right)^2\right] + \frac{1}{2}H(2\xi_1\tau)\left(\frac{c_1}{\xi_1}\right)^2$$

For $J = 2$

$$R(\infty) = \mu + \theta_1\left[\frac{c_1}{\xi_1} + \frac{c_2\rho}{\xi_2}\right] + \theta_2\left[\frac{c_2\sqrt{1-\rho^2}}{\xi_2}\right] - \frac{1}{2}\left[\frac{c_1^2}{\xi_1^2} + \frac{2c_1c_2\rho}{\xi_1\xi_2} + \frac{c_2^2}{\xi_2^2}\right]$$

$$\begin{aligned} w(\tau) = {} & H(\xi_1\tau)\left[\theta_1\frac{c_1}{\xi_1} - \left(\frac{c_1^2}{\xi_1^2} + \frac{c_1c_2\rho}{\xi_1\xi_2}\right)\right] \\ & + H(\xi_2\tau)\left[\theta_1\frac{c_2\rho}{\xi_2} + \theta_2\frac{c_2\sqrt{1-\rho^2}}{\xi_2} - \left(\frac{c_1c_2\rho}{\xi_1\xi_2} + \frac{c_2^2}{\xi_2^2}\right)\right] \\ & + \frac{1}{2}\left[H(2\xi_1\tau)\left(\frac{c_1^2}{\xi_1^2}\right) + 2H((\xi_1+\xi_2)\tau)\left(\frac{c_1c_2\rho}{\xi_1\xi_2}\right) + H(2\xi_2\tau)\left(\frac{c_2^2}{\xi_2^2}\right)\right] \end{aligned} \tag{5.4-8}$$

See Babbs and Nowman (1997) equation (23) for the general case for J state variables and ($Q \leq J$) sources of independent risk. The parameters θ_j are the market price of risk processes.

The state space form for the Kalman filter.

As discussed in Section 5.3.1 the model is put into state-space form comprising of the measurement equation (5.3-1) and the transition equation (5.3-3). In the Babbs and Nowman model, the measurements are given by N observed interest rates, for each t_k , k = 1,...,n. Denote $R(t+\tau,t)$ as the interest rate given by a bond at time t, expiring at time $t+\tau$ then the theoretical yield curve is given by :

$$R(t+\tau_i,t) \equiv -\log B(t+\tau_i,t)/\tau = A_0(\tau_i) - A_1(\tau_i)' X(t) \qquad i = 1 \ldots N \tag{5.4-9}$$

where $A_0(\tau_i) = R(\infty) - w(\tau_i)$ and $A_1(\tau_i) = H(\xi_j \tau_i)$ is a $J \times 1$ vector. Thus, in terms of (5.3-1) the ith rows of the vector $d(\Psi)$ and $Z(\Psi)$ are given by $A_0(\tau_i;\, \Psi)$ and $-A_1(\tau_i;\, \Psi)'$ respectively. The variance covariance matrix of measurement errors is assumed to have no off diagonal elements, that is having $H = h_1,\ldots,h_N$ along the diagonal.

The transition equation is given by the exact discrete time distribution from the solution to (5.4-3). The particular form of Babbs and Nowman eliminates the vector $c(\Psi)$[8] from (5.3-3) such that the transition equation has the form :

$$X_k = T(\Psi) X_{k-1} + \eta_k \tag{5.4-10}$$

where

$$T(\Psi) = \exp\left[-\kappa(t_k - t_{k-1})\right] \tag{5.4-11}$$

$$Q(\Psi) = \mathrm{Cov}(\eta_k) = \int_{t_{k-1}}^{t_k} e^{-\kappa(t_k - v)} \sigma\rho\sigma e^{-\kappa'(t_k - v)} dv \tag{5.4-12}$$

The matrices κ, σ, ρ are all of the order $J \times J$, κ being the diagonal matrix of reversion coefficients, σ, the diagonal matrix of diffusion coefficients and ρ, the correlation matrix for the Brownian motions W_j. The formal definition for the matrix exponential function, $\exp(At)$, is given by (see Golub and Van Loan (1983) for computational techniques):

$$\exp(At) = \sum_{k=0}^{\infty} \frac{1}{k!} (At)^k \tag{5.4-13}$$

In the case of interest above, where the matrix for which the exponential required is diagonal, the solution can be found explicitly. As such the matrices for T and Q evaluate to :

[8] See Lund (1997) for the general case of the discrete time distribution for the dynamics of a Gaussian term structure model.

For $J = 1$

$$T(\Psi) = e^{-\xi_1 (t_k - t_{k-1})} \qquad (5.4\text{-}14)$$
$$Q(\Psi) = \frac{c_1^2}{2\xi_1}\left(1 - e^{-2\xi_1 (t_k - t_{k-1})}\right)$$

For $J = 2$

$$T(\Psi) = \begin{bmatrix} e^{-\xi_1 (t_k - t_{k-1})} & 0 \\ 0 & e^{-\xi_2 (t_k - t_{k-1})} \end{bmatrix} \qquad (5.4\text{-}15)$$
$$Q(\Psi) = \begin{bmatrix} \frac{c_1^2}{2\xi_1}\left(1 - e^{-2\xi_1 (t_k - t_{k-1})}\right) & \frac{c_1 \rho c_2}{(\xi_1 + \xi_2)}\left(1 - e^{-(\xi_1+\xi_2)(t_k - t_{k-1})}\right) \\ \frac{c_1 \rho c_2}{(\xi_1 + \xi_2)}\left(1 - e^{-(\xi_1+\xi_2)(t_k - t_{k-1})}\right) & \frac{c_2^2}{2\xi_2}\left(1 - e^{-2\xi_2 (t_k - t_{k-1})}\right) \end{bmatrix}$$

Having put the model into state space form required for the Kalman filter, the filter can be run to recover the state variables and the parameter estimates gained via maximisation of the likelihood function (5.3-11).

Stability of the conventional KF under the Babbs and Nowman model.

As discussed in Section 5.3.2 the conventional KF can fail if the system matrices of the state-space formulation are pre-disposed for it to do so. When the parameters of the system are known, then some assurance can be given about the stability of the system of the efficient operation of the filter. When they are unknown, and it is wished to estimate them, by a numerical search procedure, such as maximum likelihood, something can be said about the regions within which they should be constrained.

For the conventional Kalman filter it is required that the absolute value of all the eigenvalues of the transition matrix T, be inside the unit circle. Looking at the form of T in (5.4-14) and (5.4-15) above it can be seen that this constraint amounts to the requirement that the reversion coefficients ξ_j are constrained to be positive. Indeed, it makes economic sense in this model that the system is constrained to be in the stable region. Secondly, it is also of concern to ensure that H is as well conditioned as possible. Where H is diagonal, the elements of which being all of the same order of magnitude, then H will be well conditioned. This is likely to be the case of the optimal solution. However, it is feasible that during numerical search the values of H can be chosen such that it will be poorly conditioned, causing problems for the inversion of F_t.

5.4.3 Investigating the case where the exact solution to the bond pricing equation is unknown

If the exact solution to the bond pricing equation (5.4-5) is not available, in order to utilise the observations along the term structure as measurements for the Kalman filter, some approximation to (5.4-5) has to be found. From Appendix 4-1, it was seen that an approximation to the bond pricing equation exists, valid for the short end of the yield curve. Therefore, the expression for the term structure in (5.4-9) can be approximated by :

For $J = 1$

$$R(t+\tau,t) \approx \mu - X_1(t) + \frac{1}{2}[\theta_1 c_1 + \xi_1 X_1(t)]\tau + \frac{1}{6}\left[-\xi_1\theta_1 c_1 - c_1^2 - \xi_1^2 X_1(t)\right]\tau^2 \tag{5.4-16}$$

For $J = 2$

$$\begin{aligned} R(t+\tau,t) \approx{}& \mu - X_1(t) - X_2(t) + \frac{1}{2}\left[\xi_1 X_1(t) + \xi_2 X_2(t) + \theta_1 c_1 + \theta_1 c_2\rho + \theta_2 c_2\sqrt{1-\rho^2}\right]\tau \\ &+ \frac{1}{6}\left[-\xi_2\theta_1 c_2\rho - \xi_2\theta_2 c_2\sqrt{1-\rho^2} - 2c_1c_2\rho - \xi_1\theta_1 c_1 - c_1^2 - c_2^2 - \xi_1^2 X_1(t) - \xi_2^2 X_2(t)\right]\tau^2 \end{aligned} \tag{5.4-17}$$

This same approximation can be reached by applying a third order exponential expansion to the function $H(x)$, (5.4-6), and substituting this into the expressions (5.4-7), (5.4-8). Thus the approximated form for the function $H(x)$ is :

$$H(x) \approx 1 - \frac{x}{2} + \frac{x^2}{6} \tag{5.4-18}$$

As further noted in Appendix 4-1, when restricted to using such an approximation as (5.4-16), (5.4-17), that is only valid for the short end of the yield curve, it may be useful to extract further information in the form of the derivatives of the term structure with respect to τ. In this manner, expressions may be found for the slope and curvature of the term structure. Taking derivatives with respect to τ, of the approximated expressions (5.4-16), (5.4-17):

For $J = 1$

$$\frac{\partial R(t+\tau,t)}{\partial \tau} \approx \frac{1}{2}\left[\theta_1 c_1 + \xi_1 X_1(t)\right] + \frac{1}{3}\left[-\xi_1\theta_1 c_1 - c_1^2 - \xi_1^2 X_1(t)\right]\tau \qquad (5.4\text{-}19)$$

$$\frac{\partial R^2(t+\tau,t)}{\partial \tau^2} \approx \frac{1}{3}\left[-\xi_1\theta_1 c_1 - c_1^2 - \xi_1^2 X_1(t)\right]$$

For $J = 2$

$$\frac{\partial R(t+\tau,t)}{\partial \tau} \approx +\frac{1}{2}\left[\xi_1 X_1(t) + \xi_2 X_2(t) + \theta_1 c_1 + \theta_1 c_2 \rho + \theta_2 c_2\sqrt{1-\rho^2}\right]$$

$$+\frac{1}{3}\left[-\xi_2\theta_1 c_2\rho - \xi_2\theta_2 c_2\sqrt{1-\rho^2} - 2c_1c_2\rho - \xi_1\theta_1 c_1 - c_1^2 - c_2^2 - \xi_1^2 X_1(t) - \xi_2^2 X_2(t)\right]\tau$$

$$\frac{\partial R^2(t+\tau,t)}{\partial \tau^2} \approx \frac{1}{3}\left[-\xi_2\theta_1 c_2\rho - \xi_2\theta_2 c_2\sqrt{1-\rho^2} - 2c_1c_2\rho - \xi_1\theta_1 c_1 - c_1^2 - c_2^2 - \xi_1^2 X_1(t) - \xi_2^2 X_2(t)\right]$$

(5.4-20)

In practice, if it is wished to ally these expressions with empirical measurements, a method may be used, such as fitting a Lagrangian polynomial to the short end of the term structure as detailed in Appendix 4-2. The derivatives of the expression for the Lagrangian polynomial then provide empirical estimates of the slope and curvature.

Comparison of Exact/Approximated term structure estimates from a simulated process.

To investigate how well the expressions (5.4-16), (5.4-17), (5.4-19), (5.4-20) may perform in describing the term structure, it is useful to quantify their performance based upon simulated data. If the approximations can be shown to be good estimates of their theoretical equivalents, then this provides justification for their use.

In their paper, Babbs and Nowman (1997) estimate both one and two factor versions of the model, using US data consisting of 507 weekly observations over the period April 1987 to December 1996. Eight maturities are used, these being 3mth, 6mth, 1yr, 2yr, 3yr, 5yr, 7yr, 10yr.

Here, a one factor generalised Vasicek process is simulated of the form (5.4-3) using the exact discretisation implied by (5.4-14). The implied instantaneous short rate is then given by μ-$X_1(t)$ and the theoretical term structure by (5.4-9) and (5.4-7). The parameters used to simulate the process are the same as those found by Babbs and Nowman for their one factor model, and reproduced in Table 5-1 below. The number of time steps used is 507, corresponding to 9.75 years of weekly data. The simulated process and resulting theoretical term structure can then be thought to be realistic.

TABLE 5-1: PARAMETERS USED TO SIMULATE ONE FACTOR B&N PROCESS

ξ_1	0.1908
c_1	0.0132
θ_1	0.6483
μ	0.0594
t_k-t_{k-1}	1/52
$R(0,0)$	0.063947
$R(\infty)$	0.101858

A simulated path for the instantaneous short rate and the implied theoretical term structure for τ=0.25, 0.5, 1yr are shown in Figure 5-1 and Figure 5-2 below.

For the short end of the yield curve, the approximation to the term structure (5.4-16) provides a good estimate for its theoretical counterpart. This may be quantified by assessing the mean squared error of the approximated term structure against the theoretical benchmark over the sample path. To fully demonstrate this, a Monte-Carlo procedure is run to obtain accurate estimates of these mean squared errors. Table 5-2 below gives Monte-Carlo estimates for this mean squared error based upon 100 sample paths for the eight maturites of interest. Antithetic variables are used in the Monte-Carlo procedure. The results given in this section must be viewed in the knowledge that they are conditional upon the values from Table 5-1.

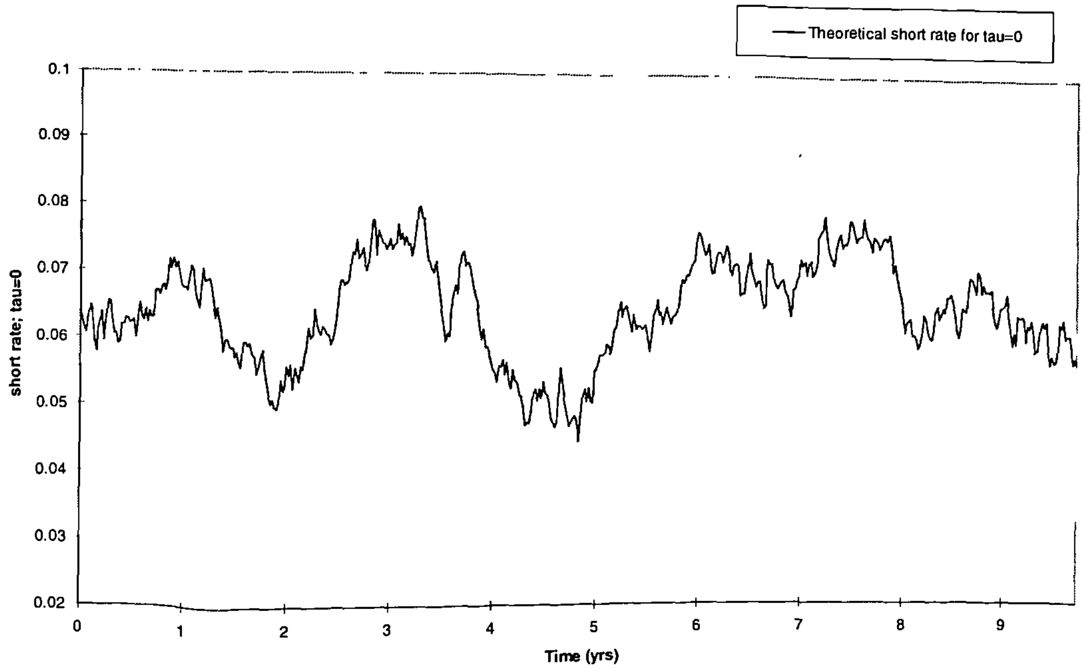


FIGURE 5-1: SIMULATED B&N PROCESS. THEORETICAL INSTANTANEOUS SHORT RATE SAMPLE PATH (TAU = 0 YRS)

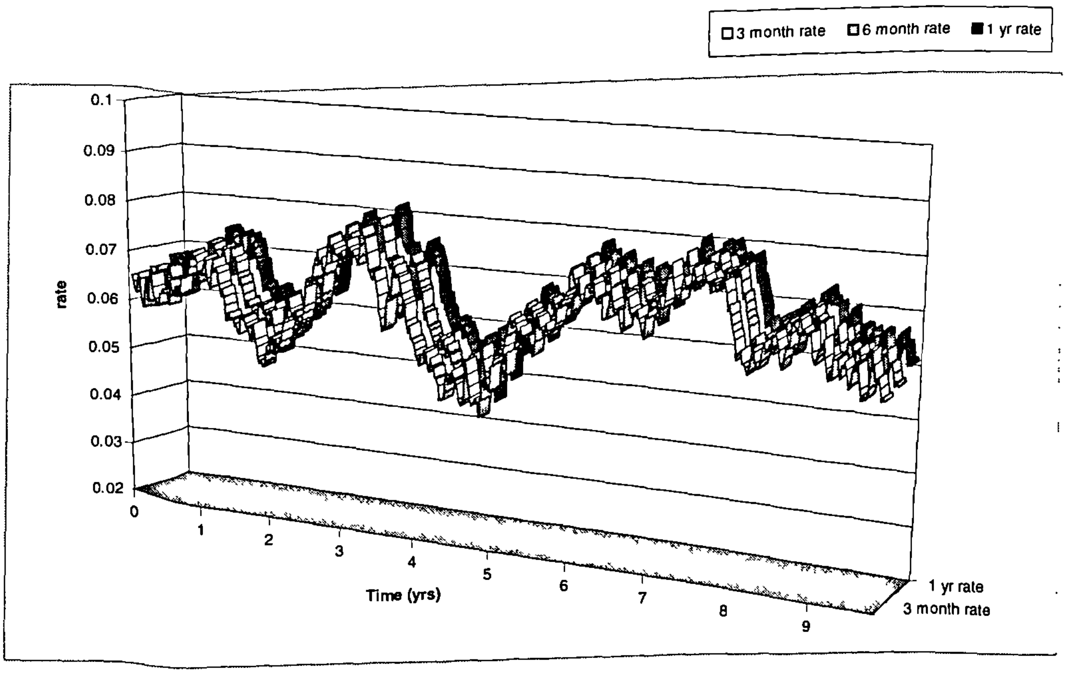


FIGURE 5-2: SIMULATED B&N PROCESS. THEORETICAL 3, 6 MONTH AND 1 YR RATE IMPLIED BY SAMPLE PATH IN FIGURE 5-1

TABLE 5-2: MONTE-CARLO ESTIMATES FOR THE MSE BETWEEN THE THEORETICAL AND APPROXIMATED TERM STRUCTURE

	Theoretical T.S. vs. Approximated T.S		Average Deviation	Mean value of theoretical rate		% Deviation
τ	MSE	S.E.	$(MSE)^{1/2}$	Mean	S.E.	
0.25	7.22E-14	(3.27E-15)	2.69E-07	0.062468	(0.001618)	4.3E-06
0.5	4.51E-12	(2.05E-13)	2.12E-06	0.063436	(0.001580)	3.35E-05
1	2.75E-10	(1.26E-11)	1.66E-05	0.065274	(0.001508)	0.000254
2	1.59E-08	(7E-10)	1.26E-04	0.068594	(0.001145)	0.001836
3	1.65E-07	(7.35E-09)	4.06E-04	0.071499	(0.001049)	0.005683
5	2.99E-06	(1.35E-07)	1.73E-03	0.076301	(0.000888)	0.022651
7	1.94E-05	(8.87E-07)	4.40E-03	0.080055	(0.000746)	0.054976
10	1.33E-04	(6.16E-06)	1.15E-02	0.084284	(0.000604)	0.136597

It is clear from Table 5-2 that the MSE grows as the time to maturity (τ) increases. This is to be expected given the nature of the approximation. The fourth column gives the Monte-Carlo estimate of the mean value for the theoretical rate. The last column provides an estimate of the percentage deviation implied by the MSE in using the approximated term structure formula. This is calculated as the ratio of $(MSE)^{1/2}$ to the estimate of the mean value for the theoretical rate. For the one year rate, it can be expected that the approximation will be valid to three decimal places. Use of the approximation at the one year rate implies an average error of 0.0000166 in approximating a data point, or expressing this as relative to the Monte-Carlo mean, 0.000254%.

As noted earlier, when restricted to using measurements only at the short end of the yield curve, it may be useful to extract extra information in the form of the slope and curvature of the term structure. The exact expressions for the slope and curvature of the term structure may be found by differentiating (5.4-9) with respect to τ. The expression for the slope is then given by (5.4-21) and the curvature by (5.4-22) below, where $w(\tau)$ is as in (5.4-7) above.

$$\frac{\partial R(t+\tau_i,t)}{\partial \tau_i} = -\frac{\partial w(\tau_i)}{\partial \tau_i} - \frac{\partial H(\xi_1 \tau_i)}{\partial \tau_i} X_1(t) \qquad (5.4\text{-}21)$$

$$\frac{\partial^2 R(t+\tau_i,t)}{\partial \tau_i^2} = -\frac{\partial w^2(\tau_i)}{\partial \tau_i^2} - \frac{\partial H^2(\xi_1 \tau_i)}{\partial \tau_i^2} X_1(t) \tag{5.4-22}$$

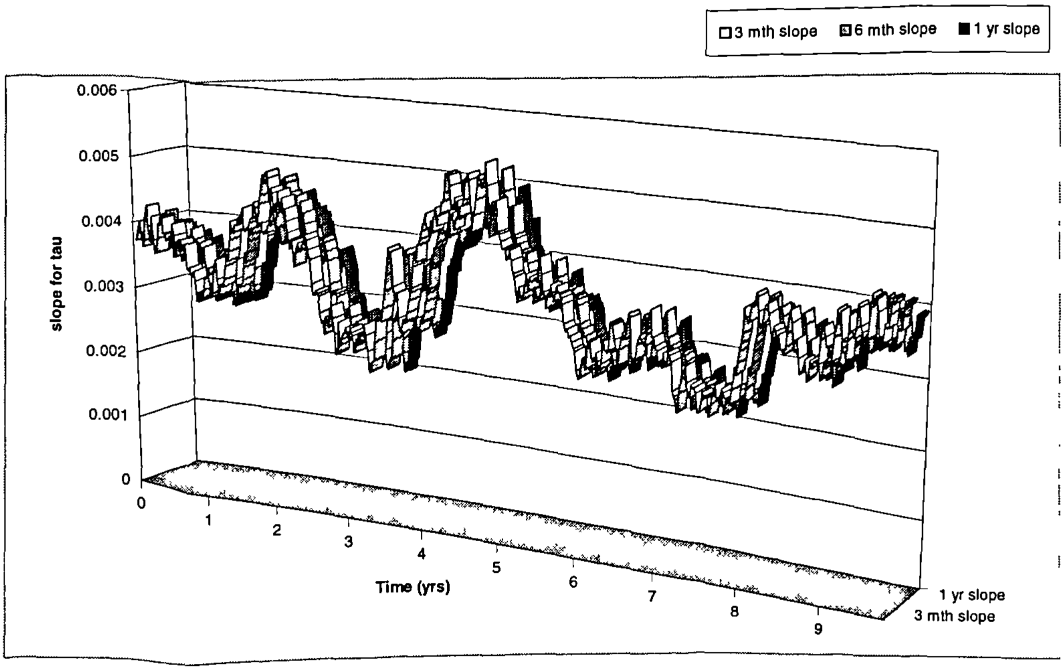


FIGURE 5-3: THEORETICAL SLOPE OF TERM STRUCTURE AT TAU = 0.25, 0.5, 1 YRS IMPLIED BY SAMPLE PATH IN FIGURE 5-1

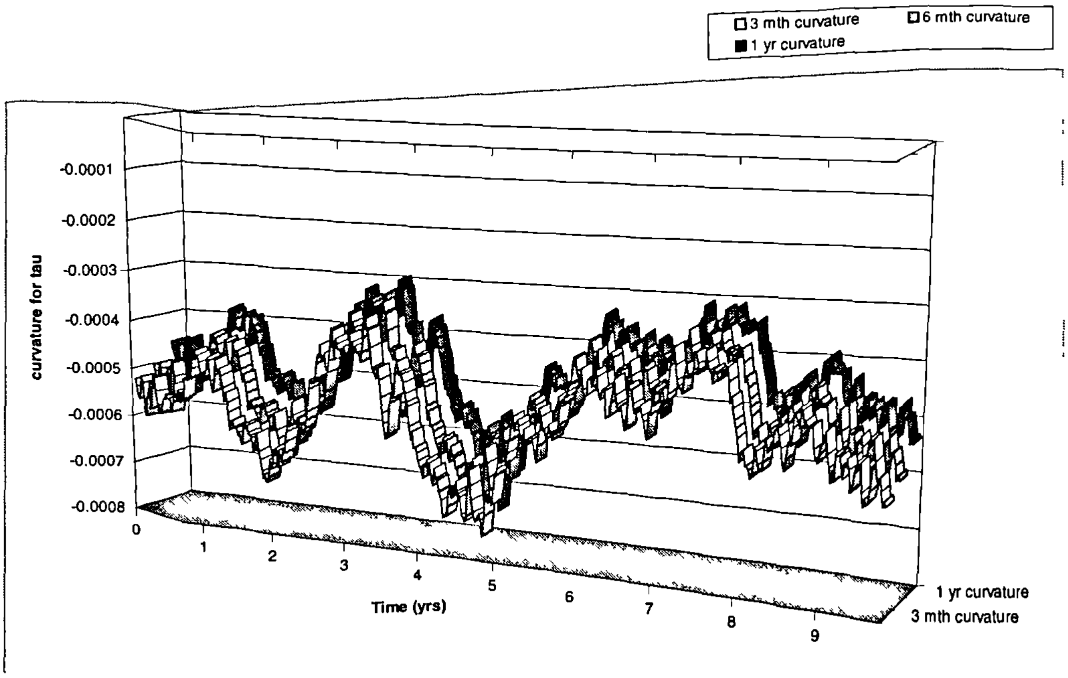


FIGURE 5-4: THEORETICAL CURVATURE OF TERM STRUCTURE AT TAU = 0.25, 0.5, 1 YRS IMPLIED BY SAMPLE PATH IN FIGURE 5-1

Figure 5-3 and Figure 5-4 above give the theoretical slope and curvature implied by the short rate sample path in Figure 5-1.

Again, if it is required to use the approximations (5.4-19) in place of (5.4-21), (5.4-22) then it is useful to quantify the degree of error involved in so doing. As above, a Monte-Carlo analysis is run to provide estimates of the error involved in approximating the slope and curvature of the term structure. Table 5-3 below gives the Monte-Carlo estimates from 100 sample paths for MSE between the theoretical slope and curvature and their approximated counterparts. Only estimates for time to maturity up to one year are presented as the deterioration in the approximation after this point renders them unusable.

TABLE 5-3: MONTE CARLO ESTIMATES FOR THE MSE BETWEEN THE THEORETICAL AND APPROXIMATED SLOPE AND CURVATURE

	Theoretical slope vs Approximated slope		Average Deviation	Mean value of theoretical slope		% Deviation
τ	MSE	S.E.	$(MSE)^{1/2}$	Mean	S.E.	
0.25	1.02E-11	(3.88E-13)	3.2E-06	0.003941	(0.000108)	0.000811
0.5	1.58E-10	(6.02E-12)	1.26E-05	0.003805	(0.000105)	0.003307
1	2.38E-09	(9.1E-11)	4.87E-05	0.003551	(9.82E-05)	0.013727
	Theoretical c'ture vs Approximated c'ture		Average Deviation	Mean value of theoretical curvature		% Deviation
τ	MSE	S.E.	$(MSE)^{1/2}$	Mean	S.E.	
0.25	6.34E-10	(2.55E-11)	2.52E-05	-0.00055	(1.51E-05)	0.045512
0.5	2.41E-09	(9.78E-11)	4.91E-05	-0.00053	(1.46E-05)	0.092678
1	8.77E-09	(3.59E-10)	9.37E-05	-0.00049	(1.36E-05)	0.192157

The results in Table 5-3 show a lessening in the accuracy with which the approximation fits the theoretical benchmark, both as time to maturity increases and as further derivatives are taken. As can be seen from (5.4-19) the approximated slope is a function of order one in τ, the curvature being independent of τ. It is obviously going to be the case that an approximation independent of τ will poorly fit a function including powers of τ up to the order e^{τ}/τ, as is the case for the theoretical curvature. Using the approximation to calculate the curvature for the one year rate implies an average error of nineteen percent. It follows that use of the approximation to describe the curvature of the term structure should be done with care, if at all.

In practice it is necessary to make empirical measurements of the slope and curvature of the term structure. This may be done by fitting a quadratic function in the form of a Lagrangian polynomial to three points on the term structure. This may then be considered consistent with the approximated form proposed for the term structure, and used above. Both the approximated theoretical form for the curvature and the Lagrangian empirical measurement of the curvature will be independent of time to maturity. The expressions detailing the Lagrangian polynomial fit to the term structure can be seen in Appendix 4-2.

It is of interest how well the quadratic Lagrangian polynomial provides estimates of the slope and curvature of the term structure. To quantify this, the Lagrangian estimates may be compared with their theoretical equivalents gained from the expression for the term structure. A Monte-Carlo procedure is again employed to estimate the MSE involved in making the measurements of the slope and curvature through a quadratic polynomial fit to the term structure against the theoretical benchmarks. The Lagrangian polynomial is fitted to the first three rates on the term structure, for τ=0.25, 0.5, 1 yrs. These values are tabulated in Table 5-4 below.

TABLE 5-4: MONTE-CARLO ESTIMATES OF THE MSE FOR THE LAGRANGIAN ESTIMATES OF SLOPE AND CURVATURE AGAINST THEORETICAL VALUES

	Lagrangian slope vs Theoretical slope		Average Deviation	Mean value of Theoretical slope		% Deviation
τ	MSE	S.E.	$(MSE)^{1/2}$	Mean	S.E.	
0.25	8.79E-12	(4.08E-13)	2.96E-06	0.003941	(0.000108)	0.000752
0.5	3.81E-12	(1.77E-13)	1.95E-06	0.003805	(0.000105)	0.000513
1	3.27E-11	(1.53E-12)	5.72E-06	0.003551	(9.82E-05)	0.001610
	Lagrangian c'ture vs Theoretical c'ture		Average Deviation	Mean value of Theoretical curvature		% Deviation
τ	MSE	S.E.	$(MSE)^{1/2}$	Mean	S.E.	
0.25	1.01E-09	(4.58E-11)	3.18E-05	-0.00055	(1.51E-05)	0.057530
0.5	5.77E-11	(2.63E-12)	7.6E-06	-0.00053	(1.46E-05)	0.014328
1	1.4E-09	(6.42E-11)	3.75E-05	-0.00049	(1.36E-05)	0.076838

From Table 5-4 it can be seen that the Lagrangian polynomial provides an acceptable fit to the theoretical benchmark in the case of the slope. The lowest percentage error involved is in the estimate of the slope for the six month rate. This is intuitive in that the six month rate is the middle of the three values to which the polynomial is fitted. The best fit for the Lagrangian curvature against its theoretical benchmark is also at the six month rate. Here, an average error of one percent against the theoretical benchmark is quite acceptable.

As it is the case that the empirical Lagrangian estimates of the slope and curvature differ from the theoretical values, Monte-Carlo estimates of the expected standard deviations of the measurement errors can be made. Under the hypothesis that the theoretical expression for the term structure completely describes the real process, the Monte-Carlo estimates of the standard deviations of the measurement errors yield base estimates for these parameters as estimated by the Kalman filter procedure. Table 5-5 below gives these Monte-Carlo estimates.

TABLE 5-5: MONTE-CARLO ESTIMATES FOR THE EXPECTED STANDARD DEVIATIONS OF THE MEASUREMENT ERRORS ON THE SLOPE AND CURVATURE

	Error on Theoretical Slope vs Lagrangian Slope			Error on Theoretical Curvature vs Lagrangian Curvature		
τ	Variance	S.E.	S.D.	Variance	S.E.	S.D.
0.25	4.37E-13	(3.19E-14)	6.61E-07	6.04E-11	(4.68E-12)	7.77E-06
0.5	1.93E-13	(1.39E-14)	4.36E-07	3.51E-12	(2.72E-13)	1.87E-06
1	1.65E-12	(1.21E-13)	1.28E-06	8.59E-11	(6.65E-12)	9.27E-06

Use of the Lagrangian slope of the term structure at the six month rate as an empirical measurement implies a standard deviation of the measurement error of 4.36E-07 when described by the one factor theoretical generalised Vasicek process.

5.4.4 Empirical investigation into the use of approximations and derivatives of the term structure

It has been seen how approximations to the theoretical term structure, its slope and curvature can be made. This section seeks to address how these may work in the case of empirical data. For this section of the analysis, Babbs and Nowman kindly supplied the data they used in the estimation work of Babbs and Nowman (1997). As noted above, this consists of US stripped data on eight maturities for τ=0.25, 0.5, 1, 2, 3, 5, 7, 10 yrs, with 507 weekly observations over the period April 1987-December 1996.

A one factor model is put into state space form and the Kalman filter is run, using the eight maturities and the exact theoretical form of the term structure as detailed in Section 5.4.2. The parameters used for the model are the same as those found by Babbs and Nowman under maximum likelihood estimation and reported in Table 1 of their paper and also above in Table 5-1 of this chapter. The path for the state variable $X_1(t)$ is recovered allowing for the reconstruction of the estimated paths for the eight maturities via the theoretical formula for the term structure (5.4-9). Figure 5-5 below reproduces Figure 2 from Babbs and Nowman (1997).

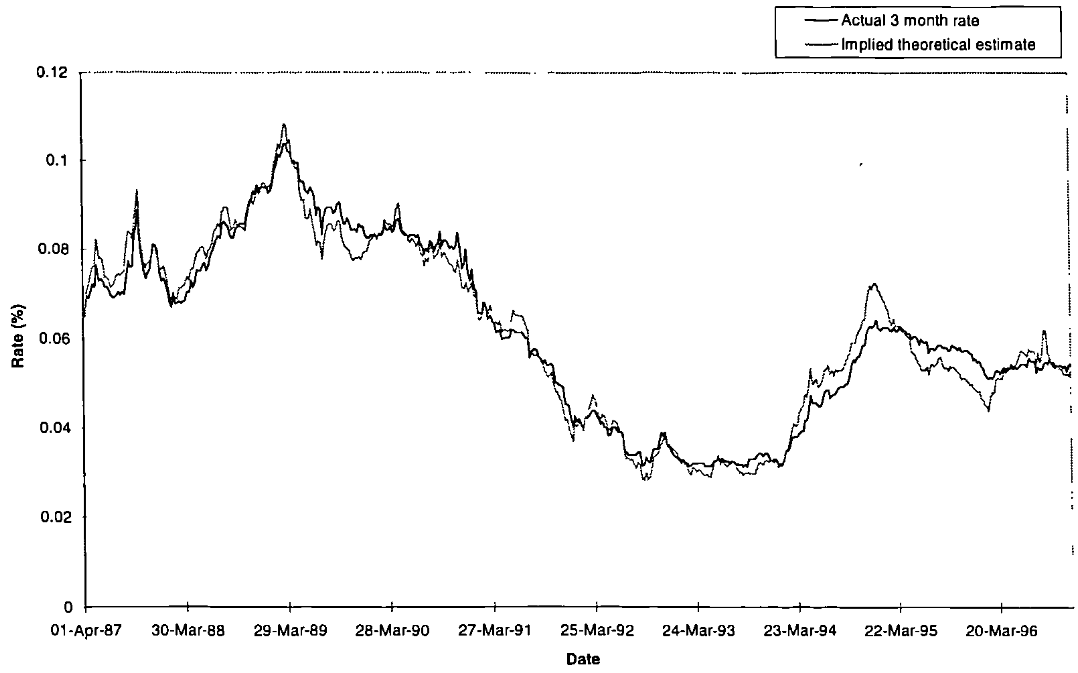


FIGURE 5-5: COMPARISON OF OBSERVED AND ESTIMATED 3 MONTH INTEREST RATE

As can be seen from Figure 5-5 the filter reproduces the path of the three month rate. The measurement error has a standard deviation of 36 basis points. Similarly, Babbs and Nowman present the fit for the six month and one year rate where the measurement errors are 22 and 4 basis points respectively. Measurements of the slope and curvature are not used, so their implied theoretical values are not to be expected to match the empirical values so well. Figure 5-6 below compares the implied theoretical slope from the filter output with the empirical slope taken from fitting a Lagrangian polynomial to the three month, six month and 1 year rates.

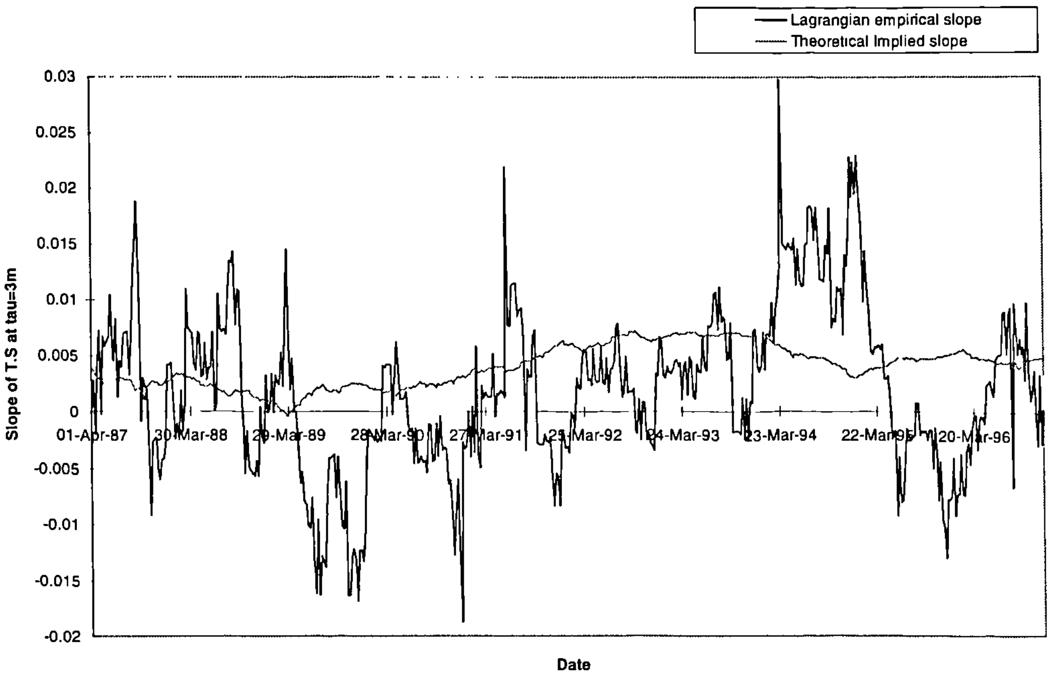


FIGURE 5-6: COMPARISON OF LAGRANGIAN EMPIRICAL SLOPE AND THEORETICAL SLOPE FOR THE TERM STRUCTURE AT TAU = 3 MONTHS

It is clear from Figure 5-6 that the theoretical slope of the term structure describes its empirical counterpart poorly. This may in part be due to a large amount of activity at the short end of the yield curve causing undue flexing, not accounted for by the one factor generalised Vaiscek process. A similarly poor relationship is visible for the six month and one year rate. The theoretical slope is constrained to be a path from the same statistical drawing as that used to create the sample path slope in Figure 5-3.

Following the suspicion that the short end of the yield curve may be responsible for the observed increased variability in the empirical slope, the slope further down the yield curve can be observed. Figure 5-7 below fits a Lagrangian polynomial to the 5 yr, 7 yr and 10 yr rate, to determine the empirical slope at the 7 yr rate.

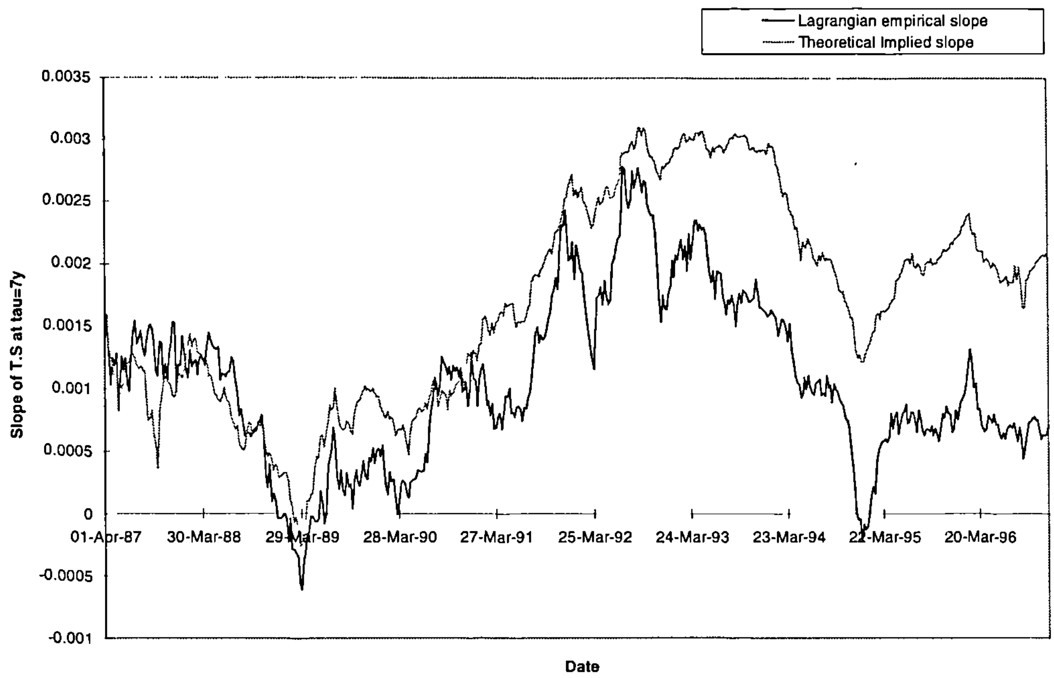


FIGURE 5-7: COMPARISON OF LAGRANGIAN EMPIRICAL SLOPE AND THEORETICAL SLOPE AT TAU = 7YRS

In Figure 5-7 above, the correlation between the empirical estimate and the theoretical slope from the Kalman filter output is improved over that observed in Figure 5-6. This adds credence to the suspicion that large amounts of activity at the short end of the yield curve cause the increased variability. It may be the case that the two factor generalised Vasicek process would be better able to capture this facet of the data. This will be approached later in the analysis. To demonstrate how the empirical slope changes as time to maturity increases Figure 5-8 below charts the empirical Lagrangian slope for the maturities 3 month through to 10 year. These are calculated by fitting a quadratic polynomial to three rates and the slope is found for the maturity at the middle of the three (except for the 0.25 year and 10 year rate where the fit is to the 0.25 year, 0.5 year, 1 year and 5 year 7 year 10 year respectively).

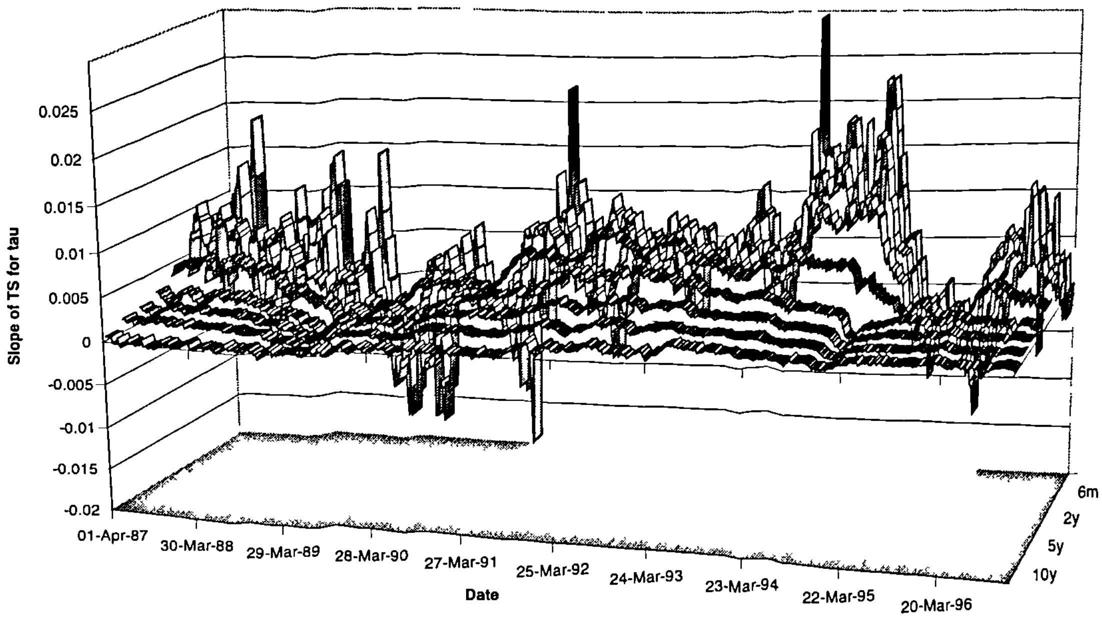


FIGURE 5-8: LAGRANGIAN EMPIRICAL SLOPE OF TERM STRUCTURE FOR MATURITIES 3 MONTH THROUGH 10 YEAR

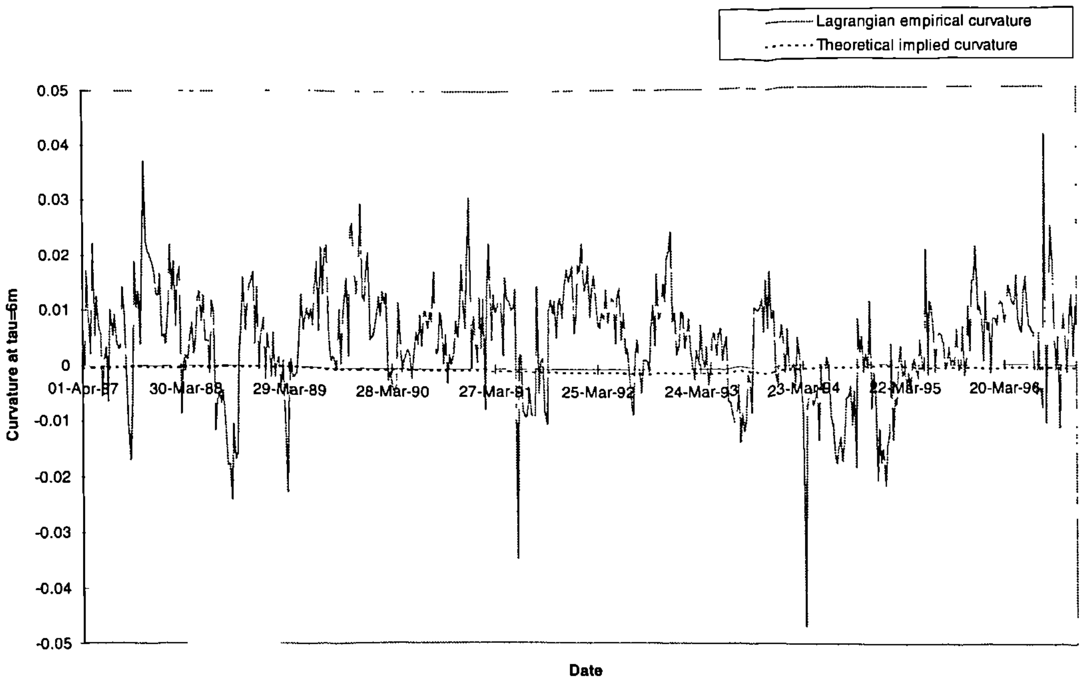


FIGURE 5-9: COMPARISON OF LAGRANGIAN EMPIRICAL CURVATURE AND THEORETICAL IMPLIED CURVATURE AT TAU = 6 MONTHS

The empirical curvature of the term structure, expectedly, also exhibits a similar departure from the values given by the theoretical term structure equation. Figure 5-9 above compares the curvature from fitting the Lagrangian polynomial to the three month, six month and one year rate with the theoretical curvature at the six month rate implied by the Kalman filter output.

It is quite obvious that the theoretical value of the curvature is a poor representation of the empirical value. The theoretical curvature is almost imperceptible in comparison with the empirical value. The theoretical curvature has the same statistical parameters as the sample in Figure 5-4. Indeed, there is little qualitative evidence in Figure 5-9 to say that they are generated by the same process. As in the case of the slope, the empirical curvature smoothes out and moves closer to its theoretical value as time to maturity increases. Figure 5-9 below shows the empirical curvature of the term structure, at times to maturity three month to ten year.

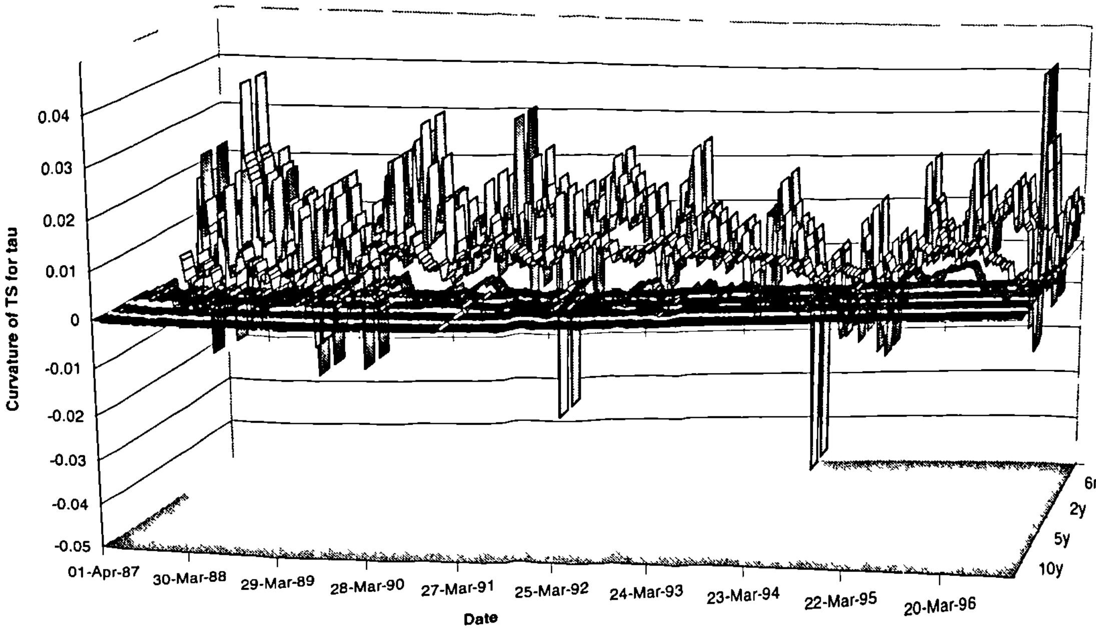


FIGURE 5-10: LAGRANGIAN EMPIRICAL CURVATURE OF TERM STRUCTURE FOR MATURITIES 3 MONTH THROUGH 10 YEAR

It is of interest to quantify the extra information that can be gained in the use of the slope and curvature as measurement data. In order to do this, these empirical measurements can be used as input to the Kalman filter. It is also of interest how well use of the approximated description of the term structure equation, its slope and curvature compare against use of the exact theoretical form. As noted above, the approximation to the term structure equation is only valid for the short end of the yield curve. It is also the case, from the empirical analysis of the data, that the theoretical (and approximated) slope and curvature describe their empirical counterparts poorly in this region of the term structure. Indeed, as is visible from Figure 5-6 and Figure 5-9, there is little qualitative evidence that they are generated by the same process. A priori, it can be expected that use of the slope and curvature, at the short end of the yield curve, as measurement data for the Kalman filter will imply large measurement errors and limited increase in information.

5.4.5 Further Kalman Filter estimates for the Babbs and Nowman model.

Here, the practical implications of the approximated form for the term structure and the use of the slope and curvature as measurements are quantified. The Kalman filter is used to uncover the hidden state variables and recover parameter estimates via maximum likelihood. The performance of the procedure using the exact theoretical form for the term structure, its slope and curvature, can then be compared against the values obtained using the approximated form.

All algorithmic implementation is made using Borland Turbo Pascal to double precision. Both the conventional Kalman filter (conventional KF) and the square root covariance filter (SRCF) as detailed in Section 5.3 are implemented. As noted in Section 5.4.2, the cases of interest here are stable under the conventional KF. The system is linear in the variables and its derivation means that the stability of the system and the economics are directly linked. The economic requirement that the impact of a piece of news on the system will always die away exponentially means that the state transition matrix will not be unstable. Indeed, no difference in the results was found by substituting the conventional KF by the SRCF algorithm. Whilst the economic specification of the system here will not alter the stability of the filtering problem, it is still prudent to use the SRCF algorithm whilst parameter optimisation is undertaken. This will help to ensure that the optimisation is not compromised by, say, instability in the inversion of the covariance matrix of innovations F_t.

Optimisation of the log-likelihood function is effected via the BFGS variant of the Davidson Fletcher Powell algorithm, from the Numerical Recipes library. Details of this method can be found in Press et al (1989) and also in Hamilton (1994). Both analytical and numerical derivatives are used in the procedure. Numerical derivatives provide much faster estimates for employment in the maximisation of the log-likelihood function. For refinement of the parameter estimates, analytical derivatives were used, although little, if any improvement was observed. Standard errors for the parameter estimates are calculated from the analytical derivatives using the procedure detailed in 5.3.3.

To ensure the proper running of the filter and the optimisation algorithm, the results as presented in Table 1 of Babbs and Nowman (1997) are reproduced. This consists of estimates for the one and two factor generalised Vasicek models, using the exact theoretical form of the term structure and all eight maturities as measurements. Table 5-6 below compares the results from the original estimation of Babbs and Nowman with the values obtained from my implementation.

TABLE 5-6: COMPARISON OF ESTIMATES FOR ONE AND TWO FACTOR MODELS AS ESTIMATED BY BABBS AND NOWMAN (1997) AND TICE (1998)

	Babbs and Nowman (1997)		Tice (1998)	
	One Factor Model	Two Factor Model	One Factor Model	Two Factor Model
ξ_1	0.1908	0.5529	0.1908	0.5529
	(0.0017)	(0.0108)	(0.0017)	(0.0042)
ξ_2		0.0652		0.0649
		(0.0120)		(0.0008)
c_1	0.0132	0.0195	0.0132	0.0193
	(0.0004)	(0.0023)	(0.00016)	(0.00019)
c_2		0.0186		0.0190
		(0.0018)		(0.00009)
ρ		-0.8360		-0.8360
		(0.0440)		(0.0044)
μ	0.0594	0.0728	0.0594	0.0730
	(0.0024)	(0.0168)	(0.0071)	(0.0417)
θ_1	0.6483	-0.0849	0.6482	-0.0849
	(0.0192)	(1.0361)	(0.1044)	(0.1770)
θ_2		0.0963		0.0963
		(1.2321)		(0.1166)
$h_1^{1/2}$	0.0036	0.0017	0.0036	0.0017
	(0.00005)	(0.00002)	(0.00004)	(0.00002)
$h_2^{1/2}$	0.0022	0.0004	0.0022	0.0004
	(0.00003)	(0.00006)	(0.00002)	(0.00002)
$h_3^{1/2}$	0.0004	0.0017	0.0004	0.0017
	(0.00008)	(0.00004)	(0.00003)	(0.00002)
$h_4^{1/2}$	0.0037	0.0028	0.0037	0.0028
	(0.00002)	(0.00004)	(0.00004)	(0.00003)
$h_5^{1/2}$	0.0042	0.0019	0.0042	0.0019
	(0.00003)	(0.000006)	(0.00004)	(0.00002)
$h_6^{1/2}$	0.0052	0.0009	0.0052	0.0009
	(0.00005)	(0.00003)	(0.00005)	(0.00001)
$h_7^{1/2}$	0.0062	0.00008	0.0062	0
	(0.00006)	(0.00007)	(0.00006)	-
$h_8^{1/2}$	0.0073	0.0008	0.0073	0.0008
	(0.00006)	(0.00003)	(0.00007)	(0.000009)
LogL	20422	24325	20422	24325
BIC	-40817	-48614	-40817	-48614

It is clear from Table 5-6 that the results from the reproduced estimations correlate very closely to those found by Babbs and Nowman. Some differences are observable in the standard error values, although this is to be expected given different implementations and machines. Babbs and Nowman (1997) present the log-likelihood values without the constant term. Here, and in all tables henceforth, the log-likelihood values are given inclusive of the constant. The BIC statistic values differ from those given in Babbs and Nowman (1997). There exist different methods for calculating the BIC statistic; see for instance Box et al (1994) or Wei (1990). The statistic used here is that given by Harvey (1989), shown in (5.4-23) below.

$$\text{BIC} = -2\log L\left(\tilde{\Psi}\right) + n\log T \tag{5.4-23}$$

As found by Babbs and Nowman (1997), the two factor model has an increased log-likelihood value over the one factor model, with the BIC statistic implying a significant increase in information content.

Having ascertained the reproducibility of the Babbs and Nowman results, eight further models are estimated, detailed in Table 5-7 below.

TABLE 5-7: MODELS ESTIMATED FOR ONE AND TWO FACTOR GENERALISED VASICEK PROCESS

Measurements		1 factor model (Table 5-8)	2 factor model (Table 5-9)
rates for tau = 0.25,0.5,1yrs	theoretical T.S approximated T.S.	model 1 model 3	model 2 model 4
rates for tau = 0.25,0.5,1yrs slope for tau = 0.5yrs curvature for tau = 0.25-1yrs	theoretical T.S approximated T.S.	model 5 model 7	model 6 model 8

The form for the theoretical and approximated term structure used, their slope and curvature are as in Section 5.4.2 and 5.4.3 above. Empirical measurements of the slope and curvature are found by fitting a Lagrangian polynomial to the three, six and twelve month rates. From the Monte Carlo analysis in Table 5-4 it is then appropriate to use the estimate of the slope at the six month rate as a measurement. Similarly, the empirical measurement of the curvature (which is independent of τ) is then allied with the theoretical expression for the curvature at the six month rate. The approximated expression for the curvature of the term structure, as shown in (5.4-19) and (5.4-20) is independent of τ. Given that the approximation to the theoretical term structure is only valid for the short end of the yield curve, it is only sensible to compare models in this region. Table 5-8 and Table 5-9 below present the estimates obtained from implementing the models in Table 5-7.

TABLE 5-8: COMPARISON OF ESTIMATES FOR ONE FACTOR GENERALISED VASICEK PROCESS USING THEORETICAL AND APPROXIMATED TERM STRUCTURE

Measurements	Rates only		Rates ,slope and curvature	
T.S. formula	Theoretical	Approximated	Theoretical	Approximated
Model	Model 1	Model 3	Model 5	Model 7
ξ_1	0.1645	0.1629	0.1577	0.1533
	(0.0055)	(0.0056)	(0.0045)	(0.0045)
c_1	0.0114	0.0114	0.0114	0.0114
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
μ	0.0728	0.0721	0.0729	0.0754
	(0.0075)	(0.0073)	(0.0076)	(0.0078)
θ_1	0.5197	0.5327	0.4815	0.4511
	(0.1110)	(0.1057)	(0.1088)	(0.1054)
$h_1^{1/2}$	0.0016	0.0016	0.0015	0.0015
	(0.00002)	(0.00002)	(0.00002)	(0.00002)
$h_2^{1/2}$	0	0	0	0
	-	-	-	-
$h_3^{1/2}$	0.0021	0.0021	0.0022	0.0022
	(0.00002)	(0.00002)	(0.00002)	(0.00002)
$h_4^{1/2}$			0.0053	0.0053
			(0.00005)	(0.00005)
$h_5^{1/2}$			0.0111	0.0111
			(0.0001)	(0.0001)
LogL	8896.7	8896.4	13311	13310
BIC	-17787	-17786	-26602	-26600

The results in Table 5-8 show that there is little to choose between the theoretical and approximated models based on the log-likelihood values. Indeed it is also the case that the parameter values found are very similar, as are their standard errors. All parameter values are realistic and comparable with those found in Table 5-6. In the knowledge that the approximated model is an approximation, the values obtained can be assessed against the theoretical model as a benchmark. For both the models using only the rates as measurements as well as the models using the rates, slope and curvature several differences between the theoretical and approximated estimates can be seen. In each case, the parameter estimate of the theoretical reversion rate is larger then that found via the approximated form. The estimates of the long term mean μ and the price of risk θ vary between the theoretical and approximated models, although not systematically so. For all the parameter estimates, the values for the approximated models are less than one half of a standard deviation away from those found by the theoretical model. The standard deviations of the measurement errors are the same between the theoretical models and the approximated models, implying that the Kalman filter estimates of the reversion coefficient, diffusion coefficient, price of risk and long term mean have absorbed the bias induced by the approximation. Standard deviations of the measurement errors for the three month, and six month rates are reduced compared to the one factor model in Table 5-6. The one factor models in Table 5-8 chooses to fit exactly to the six month rate.

Inclusion of the slope and curvature of the term structure as measurements sees the log-likelihood value increase sharply. This does not imply an increase in the performance of the model over using only the rates. The standard deviations of the measurement errors are not acceptable. Comparing these values with the Monte-Carlo estimates of the standard deviations in Table 5-5, it can be seen that only a very small part of the observed standard deviation of measurement errors for the slope and curvature can be attributable to the use of the Lagrangian empirical measurements. The standard deviation for the empirical Lagrangian slope at the six month rate is 0.0055, whereas the standard deviation for the measurement error against the implied theoretical slope found by the Kalman filter in Table 5-8 is 0.0053. The implication here is that the fit is very poor and it may be just as well to describe the slope by a fixed value at its mean. This approach would yield the same value for the standard deviation of the measurement error as for the empirical process. Similarly, the standard deviation for the empirical Lagrangian curvature is 0.009964 and the standard deviation of the measurement error in Table 5-8 is 0.0111. In this case, the implication is that it would be an improvement to describe the empirical curvature by a fixed value, rather than the theoretical curvature of the one factor generalised Vasicek process. The fact that there are only minimal changes in the parameter estimates moving from using only the rates as measurements to the inclusion of the slope and curvature adds further credence to the evidence that there is little if any benefit in using the slope and curvature as measurements here.

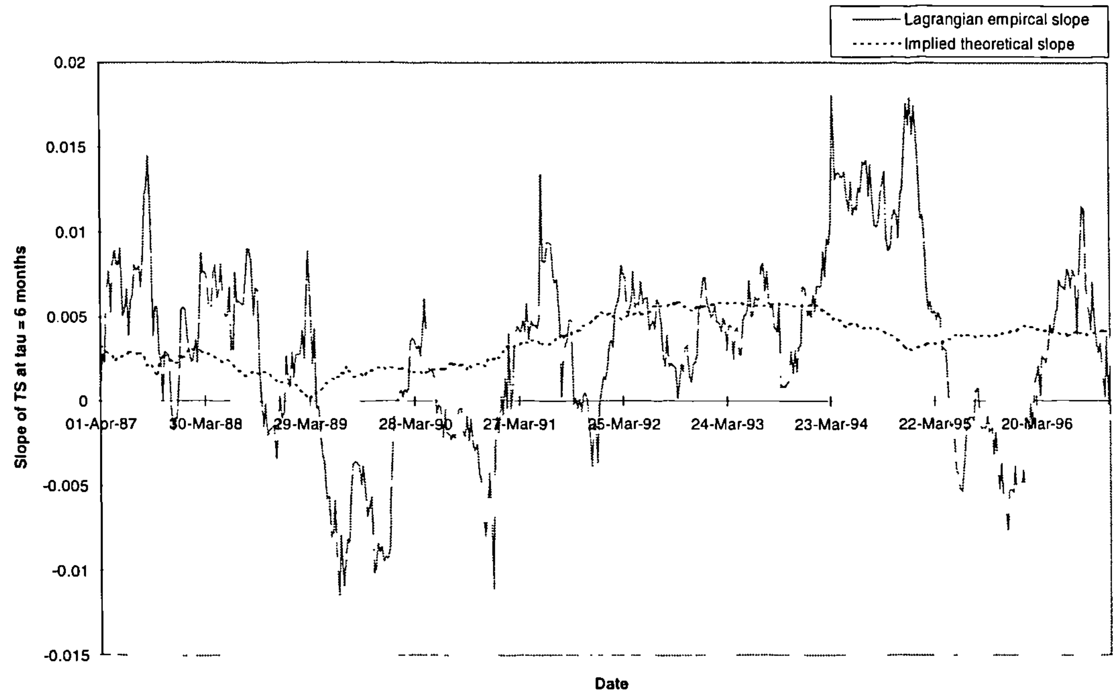


FIGURE 5-11: COMPARISON OF IMPLIED THEORETICAL SLOPE FROM ONE FACTOR ESTIMATES WITH LAGRANGIAN EMPIRICAL VALUES

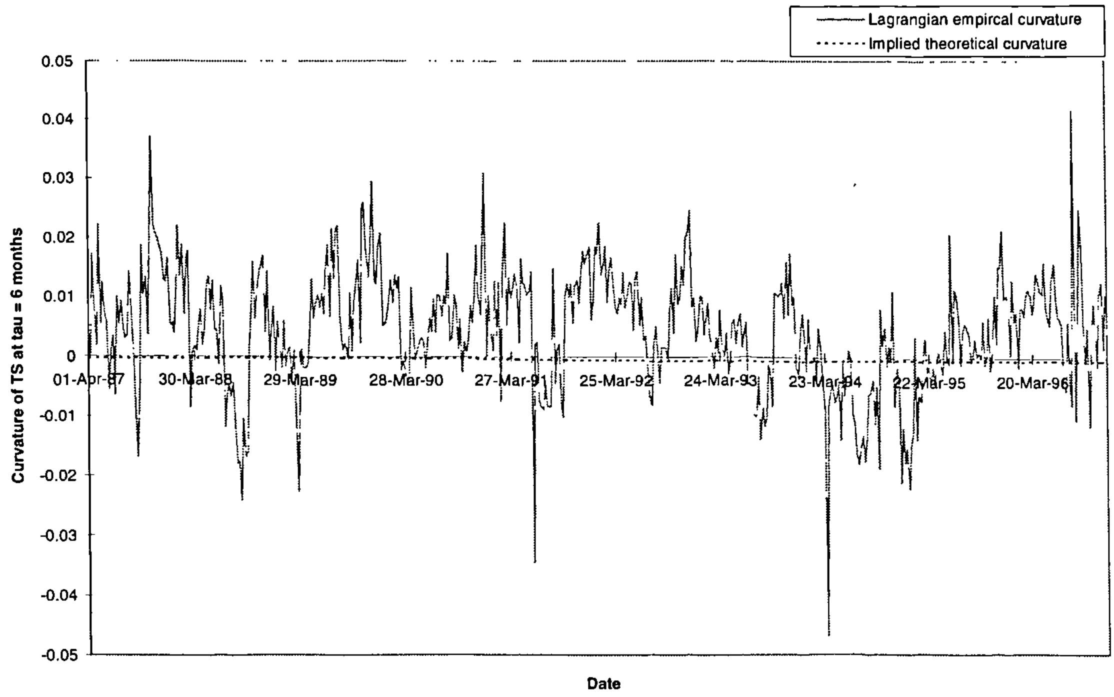


FIGURE 5-12: COMPARISON OF IMPLIED THEORETICAL CURVATURE FROM ONE FACTOR ESTIMATES WITH LAGRANGIAN EMPIRICAL VALUES

Figure 5-11 and Figure 5-12 show the fit for the Kalman filter recovered measurements of the slope and curvature implied by the estimates in Table 5-8. The correlations for the implied theoretical values against the empirical values as in Figure 5-11 and Figure 5-12 above are 0.296 for the slope and 0.035 for the curvature. This compares with correlations for the recovered measurements of the three month, six month and one year rates against their empirical values of 0.997, 1 and 0.993 respectively.

TABLE 5-9: COMPARISON OF ESTIMATES FOR TWO FACTOR GENERALISED VASICEK PROCESS USING THEORETICAL AND APPROXIMATED TERM STRUCTURE

Measurements	Rates only		Rates ,slope and curvature	
T.S. formula	Theoretical	Approximated	Theoretical	Approximated
Model	Model 2	Model 4	Model 6	Model 8
ξ_1	1.251	0.7957	1.6305	0.9721
	(0.0481)	(0.0208)	(0.0193)	(0.0096)
ξ_2	0.1272	0.0717	0.1455	0.1451
	(0.0186)	(0.0175)	(0.0070)	(0.0075)
c_1	0.0259	0.0359	0.0303	0.0484
	(0.0011)	(0.0017)	(0.00035)	(0.00096)
c_2	0.0252	0.0346	0.0266	0.0441
	(0.0011)	(0.0017)	(0.00031)	(0.00095)
ρ	-0.9124	-0.9551	-0.9177	-0.9708
	(0.0080)	(0.0044)	(0.0025)	(0.0013)
μ	0.0756	0.0764	0.0765	0.0771
	(0.0181)	(0.0343)	(0.0170)	(0.0263)
θ_1	-0.8543	-0.5897	-0.5729	-0.3390
	(0.0238)	(0.0817)	(0.1351)	(0.1028)
θ_2	0.0880	0.0483	0.1828	0.1746
	(0.0295)	(0.1030)	(0.1218)	(0.1032)
$h_1^{1/2}$	0.0002	0.0002	0	0
	(0.00002)	(0.00002)	(--)	(--)
$h_2^{1/2}$	0.0005	0.0005	0.0003	0.0002
	(0.000006)	(0.000006)	(0.000005)	(0.000005)
$h_3^{1/2}$	0.0002	0.0002	0.0008	0.0007
	(0.00002)	(0.00002)	(0.000008)	(0.000008)
$h_4^{1/2}$			0	0
			(--)	(--)
$h_5^{1/2}$			0.0080	0.0080
			(0.00008)	(0.00008)
LogL	9844	9843	15508	15527
BIC	-19663	-19661	-30986	-31024

Moving to the two factor model, the log-likelihood values increase in all cases. The BIC statistics show that compared to the one factor models, the two factor models perform significantly better. The log-likelihood values are similar for the approximated models compared to those estimated with the theoretical form. Some systematic differences in the parameter estimates are noticeable between the theoretical and approximated models. The reversion coefficients and the price of risk for the second process are larger for the theoretical models. Conversely, the diffusion coefficients, correlation coefficient, long term mean and price of risk for the first factor are larger for the approximated model. These differences do not correlate directly with those found from the estimates for the one factor model, although in both cases the reversion coefficients are larger for the theoretical models. The differences between the approximated and theoretical parameter values are more significant than those in the one factor model, particularly so for the diffusion coefficients. Again, the approximated models find standard deviations for the measurement errors which are highly similar to those of the theoretical models, implying that the bias induced in the approximation is mainly found in the other parameters.

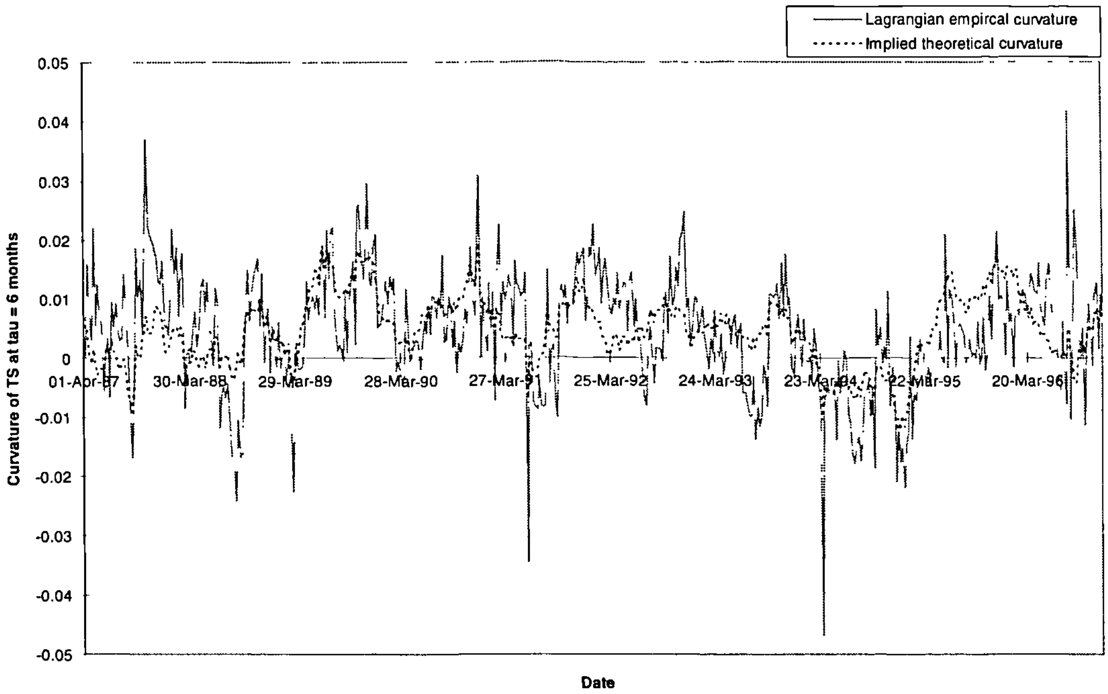


FIGURE 5-13: COMPARISON OF IMPLIED THEORETICAL SLOPE FROM TWO FACTOR ESTIMATES WITH LAGRANGIAN EMPIRICAL VALUES

The use of the slope and curvature in the two factor model provides improved measurement errors for these inputs. This is so much so that the two factor model chooses to fit exactly to the slope and finds a standard deviation of the measurement error for the curvature which is only seven percent of that for the one factor estimates. Figure 5-13 above shows the improved fit for the curvature.

The correlation coefficient for the theoretical curvature against its empirical value is 0.601, much improved over the value found for the one factor model of 0.035.

Conclusions

It is clear from this analysis that the two factor model is much better able to explain the information found in the slope and the curvature of the term structure. It may be the case that a three factor model may perform better still in its ability to fit to the curvature of the term structure. Given the large variation in the slope and curvature of the term structure as time to maturity increases visible in Figure 5-8 and Figure 5-10, it is likely that a model using measurements of the slope and curvature at both the short and long end of the yield curve may be compromised. However, the viability of using empirical Lagrangian estimates of the slope and curvature at the short end of the yield curve has been shown.

The approximated form for the term structure and its derivatives performed well in both the one and two factor cases. It appears from the estimates in Table 5-8 and Table 5-9 that the errors involved in making the approximation manifest themselves as biases in the model parameters. The standard deviations of the measurement errors remain unaffected by the approximated approach. Quantifying the biases is beyond the scope of this thesis. However, it is recognised that they exist and when assessing parameter estimates gained through such approximated forms, the standard errors should be increased to compensate.

5.5 Estimation of the dynamic mean models with the Kalman Filter

This section seeks to quantify the empirical viability of the two and three factor models developed in chapters 2 and 3 of this thesis. The models are put into state space form and estimated. In both cases, the approximated form of the term structure, as utilised in the previous section, is employed. For the two factor model, the state space form is easily found. The three factor model, is non-linear in the variables and as such cannot be directly estimated using the linear Kalman filter. A variant of the three factor model is used, to allow estimation, by assuming that the instantaneous short rate is observable.

There are two main advantages of the two and three factor dynamic mean models. The first is that they are reduced forms of a larger economic framework. As such, meaning can be ascribed to the parameter estimates with respect to their economic derivation. This is in contrast to many interest rate models where only vague economic assignation may be made to the parameters. The second advantage comes in that the models dynamics are largely driven by their deterministic form. The noise in the form of Brownian motion accounts for only a very small part of the dynamics of the system.

The disadvantages, in terms of estimation, come in the loss of analytical tractability. The exact form for the bond pricing equation and the exact discretisation of the process for the state variables are not available. This is often reason enough for many researchers to avoid estimation of a model. Here, approximations are used, which undoubtedly compromise the efficiency and introduce biases into the parameter estimates. Measures are taken to try to minimise these problems.

Four models are estimated, detailed in Table 5-10 below.

TABLE 5-10: MODELS ESTIMATED FOR TWO AND THREE FACTOR DYNAMIC MEAN PROCESSES

Measurements		2 factor model (Table 5-11)	3 factor model (Table 5-13)
rates for tau = 0.25,0.5,1yrs	approximated T.S.	model 9	model 10
rates for tau = 0.25,0.5,1yrs slope for tau = 0.5yrs curvature for tau = 0.25-1yrs	approximated T.S.	model 11	model 12

5.5.1 Estimation of the two factor model.

Recall from chapter two, the form of the two factor model.

$$\begin{aligned} dr &= \alpha(x-r)dt + \sigma_r dz_r \\ dx &= \beta\big(pr + (1-p)\mu - x\big)dt \end{aligned} \qquad (5.5\text{-}1)$$

It is assumed that there is only one source of risk, this being on the process for r. It is straightforward to put the model (5.5-1) into state space form for the Kalman filter. The exact form for the bond pricing equation and discretisation for the processes for the state variables are not available. The approximation to the term structure, as in Appendix 4-1 of chapter 4 and used for the generalised Vasicek process in 5.4.3 is employed. As such, the expression for the term structure relating to (5.5-1) can be expressed as :

$$\begin{aligned} R(t+\tau,t) \approx\; & r + \frac{\big(\alpha(x-r) - \lambda_r\sigma_r\big)}{2}\tau \\ & + \frac{1}{6}\Big[-\alpha\big(\alpha(x-r) - \lambda_r\sigma_r\big) + \alpha\Big(\beta\big(pr + (1-p)\mu - x\big)\Big) - \sigma_r^2\Big]\tau^2 \end{aligned} \qquad (5.5\text{-}2)$$

Factoring this expression in terms of the state variables x and r, yields the form for the measurement equation.

$$\begin{aligned} R(t+\tau_i) \approx\; & \left(1 - \frac{\alpha}{2}\tau_i + \frac{1}{6}\big(\alpha^2 + \alpha\beta p\big)\tau_i^2\right)r + \left(\frac{\alpha}{2}\tau_i - \frac{1}{6}\big(\alpha^2 + \alpha\beta\big)\tau_i^2\right)x \\ & - \frac{\lambda_r\sigma_r}{2}\tau_i + \frac{1}{6}\Big(\alpha\big(\lambda_r\sigma_r + \beta(1-p)\mu\big) - \sigma_r^2\Big)\tau_i^2 \end{aligned} \qquad (5.5\text{-}3)$$

The coefficients on r and x yield the values for the ith row of the matrix $Z(\Psi)$, the constant term yields the ith row of the vector $d(\Psi)$. In this analysis, and consistent with 5.4.3, measurements are used for the three, six and twelve month rates. Additional information can be used in the form of the slope and curvature, which are found from the first and second derivatives of (5.5-3) with respect to τ. The measurement errors are modelled by parameters, $h_1 \ldots h_N$ as the diagonal elements of H_t implying maturity specific error variances. The transition equation is given by the Euler discretisation of the system (5.5-1). A second order discretisation could be employed, but this would bring in terms which are non-linear in the variables, not allowing for application of the linear Kalman filter.

$$\begin{bmatrix} r_t \\ x_t \end{bmatrix} = \begin{bmatrix} 1-\alpha\Delta t & \alpha\Delta t \\ \beta p\Delta t & 1-\beta\Delta t \end{bmatrix} \begin{bmatrix} r_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \beta(1-p)\mu\Delta t \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \tag{5.5-4}$$

The error terms ε_t and η_t will comprise both the discretisation error and the modelled process noise. Modelled process noise is only implemented on the process for *r*. The Kalman filter accommodates the process noise via the covariance matrix *Q*.

To estimate the parameters for the matrix *Q*, initial estimates of the parameters of the system are found setting the diagonal elements for *Q* to a large value and optimising the system. A Monte Carlo procedure is run as follows. The exact form for the processes of the state variables may be better approximated by a numerical integration routine such as the predictor corrector method. The predicted estimates are, essentially, the Euler discretisation of the process using the corrected estimates from the previous period. The corrected minus the predicted estimates then provide an estimate of the discretisation error involved in estimating each point. 500 Monte Carlo paths are simulated and the variance for the discretisation error is found. The Monte-Carlo estimates then yield a value which can be used as a good approximation for the diagonal elements of *Q*. The variance of the error on the process for *r* is then the discretisation error plus the variance of the modelled noise, $\sigma_r^2\Delta_t$.

The data used for the estimation is the same as that in the previous section. Two models are run for the two factor case. These differ in the measurements used; the first model using the three, six and twelve month rates, the second additionally uses the slope at the six month rate and the curvature, found empirically as in the previous section. The estimation results are now presented. All estimations are effected as described in the previous section, using the SRCF filter.

TABLE 5-11: PARAMETER ESTIMATES FOR DYNAMIC MEAN TWO FACTOR MODEL

Measurements T.S Formula Model	Rates only Approximated Model 9	Rates, slope and curvature Approximated Model 11
α	0.0998	0.0614
	(0.0022)	(0.0026)
β	0.4956	0.6002
	(0.0232)	(0.0234)
μ	0.0612	0.0728
	(0.0023)	(0.0010)
λ	-0.5008	-0.3341
	(0.0898)	(0.0251)
σ	0.0113	0.0112
	(0.0001)	(0.0001)
$h_1^{1/2}$	0.0014	0.0014
	(0.000009)	(0.000007)
$h_2^{1/2}$	0	0
	--	--
$h_3^{1/2}$	0.0020	0.0019
	(0.00002)	(0.00002)
$h_4^{1/2}$		0.0055
		(0.00006)
$h_5^{1/2}$		0.0101
		(0.0001)
p	-3.3852	-3.9636
	(0.0196)	(0.0261)
LL	8997	13404
BIC	-17974	-26783

Several points may be noted from the estimates gained in Table 5-11. The value for α is less than that for β. This is contrary to the assumptions made about the parameter values in chapter 2. In point, this may only occur, if one of the IS-LM parameters, *b*, *k*, or *u* is negative. As discussed in the chapter 3, it is the possibility of *k* being negative that motivates the three factor variant of the model. This provides further justification for the use of the three factor model.

The parameters for the diffusion coefficient and the long term mean are both realistic. The diffusion coefficient is not large, in agreement with the notion that the noise serves only a minimal role in the overall dynamics of the system. The measurement errors are also realistic, although they could be considered relatively large in comparison with those obtained for the comparable estimation of the generalised Vasicek process. The value for p is negative, as postulated.

The form of the dynamics of the system is dependent upon the parameter values, as discussed in chapter 2. Here, it can be seen that the model is stable and will (in deterministic form) exhibit damped oscillatory behaviour for r and x towards μ. That is, that the condition $(\alpha-\beta)^2+4\alpha\beta p<0$ is satisfied. The deterministic dynamics of the state variables is thus found to be useful in describing the business cycle behaviour of empirical rates.

5.5.2 Estimation of the three factor model.

The three factor model (3.3-4) is not so straightforward to put into state space form. In its entirety, it has cross terms in the state variables, which preclude application of the linear Kalman filter. As a method for avoiding this, it may be assumed that the instantaneous short rate is observable. This may be done, by fitting a Lagrangian polynomial to the three, six and twelve month rates, and then extrapolating the zero month rate. The formula for this is as follows.

$$R(t+0,t) \approx 2\tfrac{2}{3}R(t+0.25,t) - 2R(t+0.5,t) + \tfrac{1}{3}R(t+1,t) \tag{5.5-5}$$

Proceeding, the approximated expression for the term structure is as in (5.5-2) above. Factoring in terms of the state variables x and p one obtains :

$$\begin{aligned} R(t+\tau_i,t) \approx {} & \left(\frac{\alpha}{2}\tau_i-\frac{1}{6}\left(\alpha^2+\alpha\beta\right)\tau_i^2\right)x+\frac{1}{6}\left(\alpha\beta(r_t-\mu)\tau_i^2\right)p \\ & +r-\frac{\left(\alpha r_t+\lambda_r\sigma_r\right)}{2}\tau_i+\frac{1}{6}\left(\alpha\left(\alpha r_t+\lambda_r\sigma_r+\beta\mu\right)-\sigma_r^2\right)\tau_i^2 \end{aligned} \tag{5.5-6}$$

The transition equation is given by the Euler discretisation of the differential equations for x and p in the system (3.3-4).

$$\begin{bmatrix} x_t \\ p_t \end{bmatrix} = \begin{bmatrix} 1-\beta\Delta t & \beta(r_{t-1}-\mu)\Delta t \\ \gamma\phi(\mu-r_{t-1})\Delta t & 1-\gamma\Delta t \end{bmatrix}\begin{bmatrix} x_{t-1} \\ p_{t-1} \end{bmatrix}+\begin{bmatrix} \beta\mu\Delta t \\ \gamma\left(\delta-\phi\mu\left(\mu-r_{t-1}\right)\right)\Delta t \end{bmatrix}+\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix}$$

(5.5-7)

As there is assumed to be no modelled noise on the processes for x and p, the error terms ε_t and η_t are solely the discretisation error. The expressions for the measurement and transition equation are now functions of the time dependent parameter r_t.

Stability of the three factor model: An example of conventional KF failure.

It is of interest if the state transition matrix in (5.5-7) is stable. This is quantifiable by the eigenvalues of the state transition matrix (STM), which for stability have to have absolute value less than one. The expression for the eigenvalues of the particular form of the STM in (5.5-7) is as follows.

$$\lambda = 1 - \frac{1}{2}\Delta t\left[\beta + \gamma \pm \sqrt{(\beta-\gamma)^2 - 4\beta\gamma\phi(r_{t-1}-\mu)^2}\right] \qquad (5.5\text{-}8)$$

These will most likely be complex, given that ϕ is large compared to β and γ and for reasonable parameter values can have absolute value greater than one if r is far from μ. The maximum of the absolute values of all the complex eigenvalues is termed the spectral radius (Edgar (1992)).

To illustrate the effect of the instability of the STM on the conventional KF, the output from the SRCF and conventional KF can be compared when the spectral radius exceeds 1. From the form of (5.5-8), it can be seen that the eigenvalues are time dependent. For illustrative purposes a sample path for the three factor system is found, such that the spectral radius starts below 1 and increases over time (as r diverges from μ). Table 5-12 details the values used to create the sample path, shown in Figure 5-14

TABLE 5-12: PARAMETER VALUES USED FOR SIMULATED PATH IN FIGURE 5-14

Lorenz Parameters		Scale Parameters		Volatility	
α	0.8	μ	0.07	σ	0.02
γ	7.89	ϕ	90000	λ	0.1
δ	23	β	8		

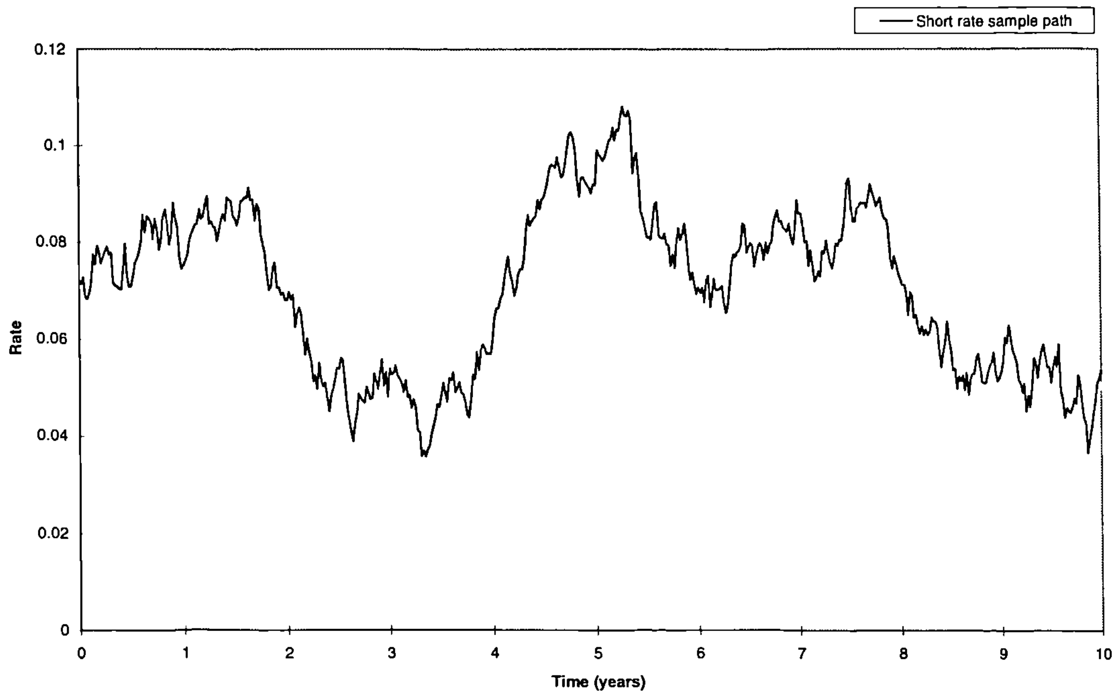


FIGURE 5-14: SAMPLE PATH FOR PARAMETER VALUES IN TABLE 5-12

The spectral radius can be charted as it changes over time as a function of $(r_t - \mu)$. This is shown in Figure 5-15 from which it can be seen that the spectral radius quickly moves above the critical value of 1. For the conventional KF this will be detrimental to the filter performance. Taking the sample path shown above, and the implied three, six and twelve month rates as input to the conventional KF, the paths for the state variables x and p are recovered. As noted in section 5.3.2 the proper running of the filter can be ascertained by ensuring that the error covariance matrix is symmetric. The error in the loss of symmetry is symptomatic of the lack of robustness of the conventional KF. For the model in question here, there are two state variables x and p and as such the error covariance matrix is of dimension 2. Figure 5-16 below shows the error between the off diagonal elements of the error covariance matrix for the conventional KF when the filter is run.

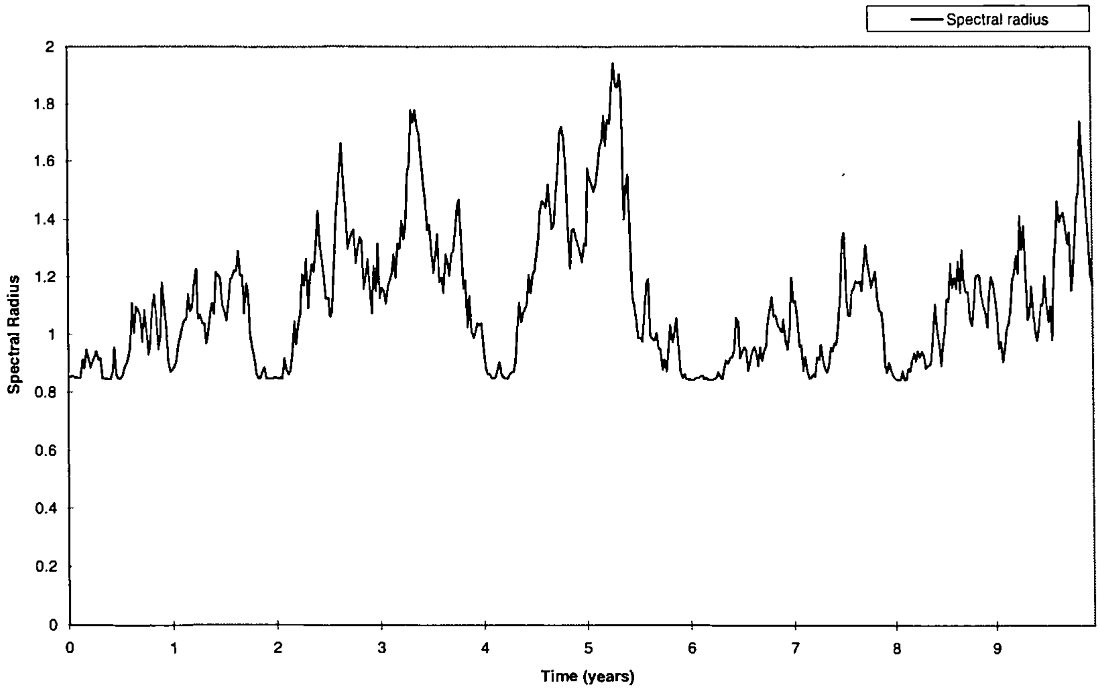


FIGURE 5-15: SPECTRAL RADIUS OF STM ASSOCIATED WITH SAMPLE PATH IN FIGURE 5-14

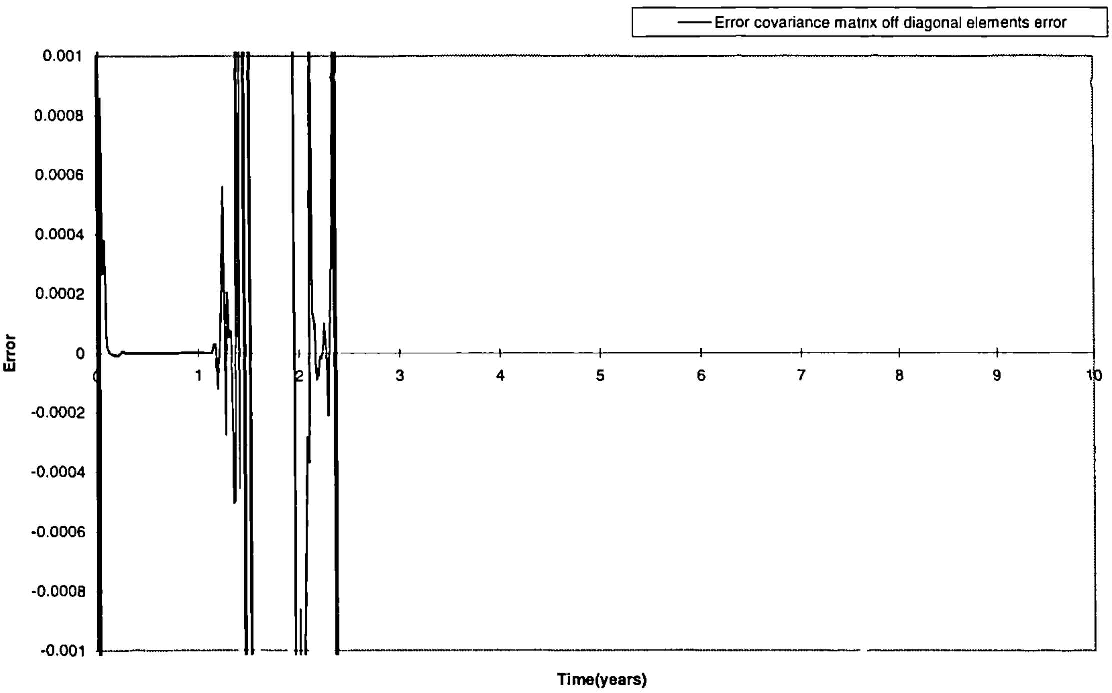


FIGURE 5-16: ERROR FOR OFF DIAGONAL ELEMENTS OF ERROR COVARIANCE MATRIX FOR CONVENTIONAL KF

From Figure 5-16 it is noticeable that the error for the first one and a half years is negligible. As the spectral radius increases, the error grows until around two and a half years when the filter fails completely. The error is transferred via the Kalman gain to the state variable estimates. Figure 5-17 shows the error on the estimate for the state variable p.

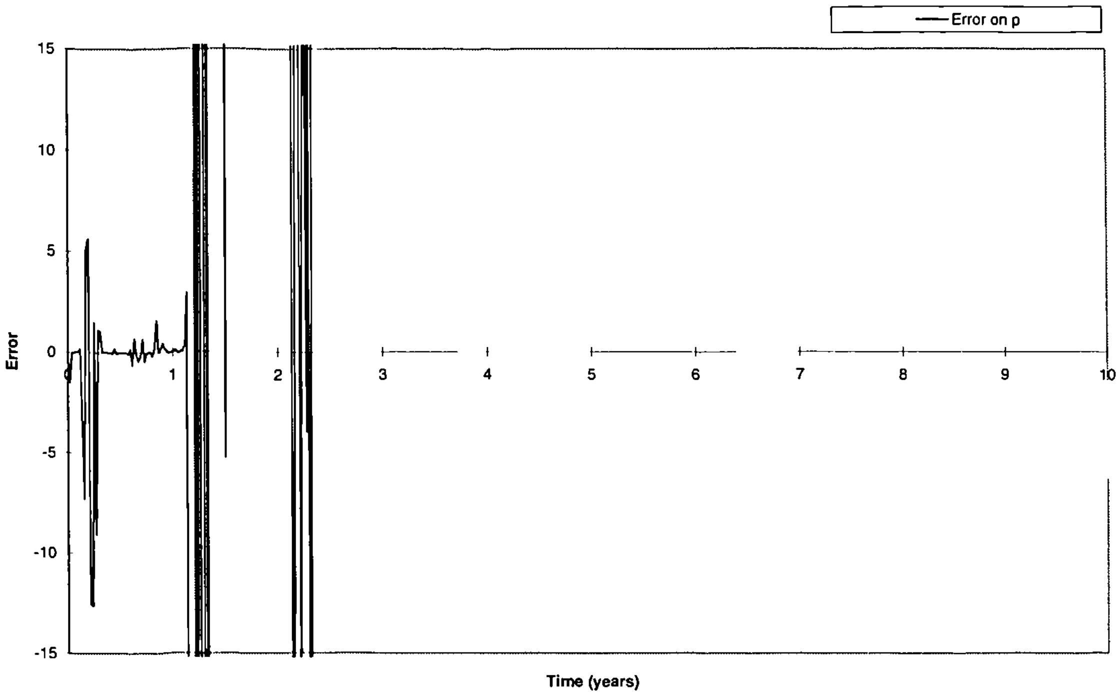


FIGURE 5-17: ERROR ON ESTIMATE FOR p (CONVENTIONAL KF)

The error on the estimate for p is correlated highly with the error on the off diagonal elements of the error covariance matrix. Conversely, the SRCF does not suffer from loss of symmetry on the error covariance matrix, as it is symmetric by construction. Applying the SRCF filter to recover the state variables, finds that the stability of the filter is not compromised by the size of the spectral radius of the STM. Figure 5-18 below shows the error on the recovered state variable p from application of the SRCF. The largest error (apart from the initial error associated with the use of a diffuse prior) is of the order of 0.01. Comparing Figure 5-18 and Figure 5-15 shows little correlation between the error and the spectral radius. The error between five and six years is associated with p fluctuating rapidly in this region, rather than the size of the spectral radius.

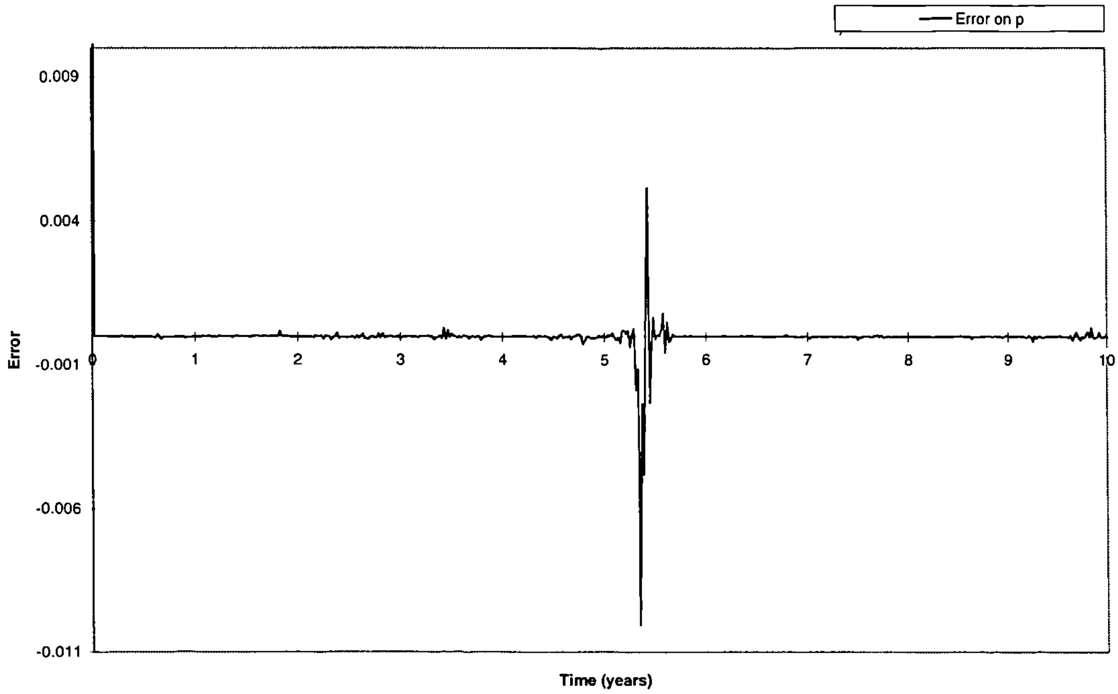


FIGURE 5-18: ERROR ON ESTIMATE FOR p (SRCF)

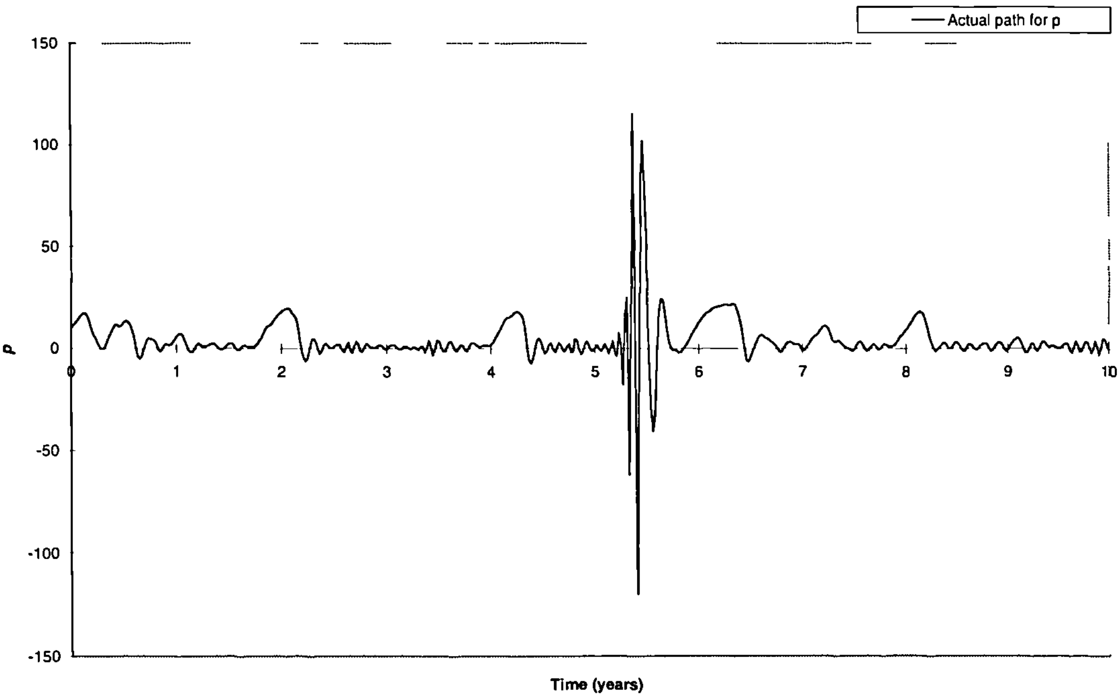


FIGURE 5-19: ACTUAL PATH FOR P

Figure 5-19 shows the actual path for p. Comparing the magnitudes of the error with the actual path for p shows the error to be of the order of 0.01%.

Estimation results for the three factor model

TABLE 5-13: PARAMETER ESTIMATES FOR DYNAMIC MEAN THREE FACTOR MODEL

Measurements T.S Formula Model	Rates only Approximated Model 10	Rates, slope and curvature Approximated Model 12
α	0.7001 (0.0202)	0.5514 (0.0132)
β	0.1650 (0.0146)	0.1623 (0.0127)
μ	0.0626 (0.0005)	0.0620 (0.0006)
λ	-0.4350 (1.6539)	-0.6463 (1.5808)
σ	0.0105 (0.0041)	0.0100 (0.0025)
$h_1^{1/2}$	0.0017 (0.00002)	0.0018 (0.00002)
$h_2^{1/2}$	0.0029 (0.00003)	0.0029 (0.00003)
$h_3^{1/2}$	0.0044 (0.00004)	0.0042 (0.00004)
$h_4^{1/2}$		0.0042 (0.00004)
$h_5^{1/2}$		0.0120 (0.0001)
γ	0.2388 (0.0225)	0.2101 (0.0227)
δ	26.275 (3.029)	21.230 (2.543)
ϕ	91031 (1115)	92659 (1211)
LL	8130	13359
BIC	-16235	-26688

The results from optimising the log-likelihood function for the three factor model are given in Table 5-13. Several things are evident from the values for the parameters found.

The values of the long term mean and the diffusion coefficient are realistic. The diffusion coefficient can be considered small. This is in-keeping with the belief that the deterministic dynamics produce the large scale fluctuations in the model. The two fixed points around which the attractor is formed when the condition for chaos is satisfied are found at $r = x = \mu \pm \sqrt{\delta - 1/\phi}$. Together, δ and ϕ determine the scale of the system, and the fixed points can be considered to be roughly half way between μ and the minimal/maximal extent of the process for r. Here the fixed points occur at $\mu \pm 0.01667$ for model 10 and $\mu \pm 0.01478$ for model 12. With reference to the empirical process in Figure 5-5 these values are credible.

For the dynamics of the system to exhibit chaos, the value for δ has to be greater than the critical value δ_H as defined in section 3.3. Here, that condition is satisfied for model 10, using only the rates as measurements, as δ and δ_H are 26.275 and 20.534 respectively. For model 12, including the slope and curvature, δ and δ_H are 21.230 and 23.695 and as such the condition for chaos is not satisfied. This may be somewhat confused by the poor fit of the model to the slope and curvature. Comparison of the BIC statistics between Table 5-13 and Table 5-11 shows that the two factor model is preferred independent of the measurements used. Overall, the large standard deviations for the measurement errors and not inconsiderable standard errors on the parameter estimates imply that the model is not a good description of the process.

There are significant reasons for believing that the estimation procedure, rather than the three factor process itself, is responsible for the poor quantitative results. Both the expression for the term structure and the transition equation are found from approximations. The transition equation, in particular, uses only a first order approximation. The instantaneous short rate is not treated properly, as an unobserved state variable, but rather a proxy is used in the form of the extrapolated rate from the polynomial fit to the three, six and twelve month rate. The implications of using proxies for the instantaneous short rate are investigated in Chapman, Long and Pearson (1998). They describe a procedure for evaluating the biases involved in the use of proxies in the form of the one and three month rate. Their results suggest that the problem is not significant for a range of affine models. However, they also investigate the non-linear functional form of Ait-Sahalia (1996), finding that significant biases result from the use of a proxy. Their findings may lead one to expect the proxy problem to be significant here.

Despite the use of spatial data, it may be expected that the time series used is not long enough to sufficiently define the existence of an attractor. Again, with reference to Figure 5-5, it may be supposed that the time series covers as little as one and a half cycles of the attractor. In this case, it would not be expected that the parameters could be recovered with a high degree of accuracy. Given that the reversion parameter estimates in Table 5-13 are small, this is consistent with the attractor evolving only slowly. These may be compared with the heuristic values of the reversion parameters and resultant short rate paths for the figures in chapter 3[9].

Conclusion

Estimation of the two and three factor dynamic mean models has been somewhat compromised by issues of tractability. Stability issues associated with the potentially chaotic three factor model have been overcome by the use of the numerically superior SRCF algorithm. However, approximations and proxies employed have most likely led to biases obfuscating the true estimates. The short time series, in terms of the number of business cycles, is a complicating issue for the three factor model, which requires these large scale dynamics to accurately locate the attractor.

[9] The values of the reversion parameters for most of the sample paths in chapter 3 are $\alpha = 5$, $\beta = 0.5$, $\gamma = 0.41667$.

Measures taken to remedy the shortcomings noted should anticipate improved results over those presented here.

5.6 Conclusion

The problem of estimating models couched in the dynamic mean framework, proposed in chapter 2, has motivated the investigation of Kalman filtering techniques. The conventional Kalman filtering algorithm is noted to suffer from numerical instabilities. In a departure from much of the finance literature, a method of the class known as 'square-root' filters is discussed and proposed as a superior alternative. A problem of particular interest in finance is the evaluation of analytical derivatives of the log-likelihood function. To be consistent with the numerically superior SRCF, I have devised a new algorithm for evaluating the analytical derivatives of the likelihood function which has several desirable numerical properties.

Application of the Kalman filter is made to one and two factor generalised Vasicek processes as described in Babbs and Nowman (1997). The use of approximated forms for the term structure is investigated and compared with theoretical equivalents. The performance of such approximations is of interest, as the exact discretisation of the bond pricing equation will not generally available for dynamic mean models which exhibit deterministic complex behaviour. The approximations are found to introduce some biases into the parameter estimates.

The dynamic mean models proposed in chapters 2 and 3 of the thesis are estimated with the SRCF optimisation. The procedure finds realistic values for the parameter estimates in both the two and three factor models. On the basis of the BIC, the two factor model is preferred. Cumulative biases, proxy problems and the short nature of the time series may adversely impact on the estimates for the three factor model.

APPENDIX 5-1: ALGORITHM FOR HOUSEHOLDER REDUCTION OF A MATRIX TO UPPER TRIANGULAR FORM

```
{Computes the Householder triangularisation of a nx(n+r) matrix. Output
in A is an upper triangular nxn matrix right adjusted in the array}
{ A specialisation of the Householder triangulisation routine for use
with the Morf-Kailath temporal/observation update algorithm as per
Grewal and Andrews pg 256, 234}
{ Amendments to reflect the practical details shown in Golub and van
Loan 1983 pgs38-42 }
{ TYPE definitions required :
       real                : double;            {optional}
       RealArrayNPbyNP     : ARRAY[1..np,1..np] OF real;
       RealArrayNP         : ARRAY[1..np] OF real;}

PROCEDURE hh_general(VAR A:RealArrayNPbyNP; n_dim,r_dim:byte);

VAR sigma : real;
    alpha : real;
    beta  : real;
    i,k,j : byte;
    v     : RealArrayNP;
    max_abs : real;


BEGIN
 FOR k := n_dim DOWNTO 1 DO BEGIN
  max_abs := 0.0;
  FOR j := 1 TO r_dim+k DO
   IF abs(A[k,j]) > max_abs THEN max_abs := abs(A[k,j]);
  sigma := 0.0;
  FOR j := 1 TO r_dim+k DO
   sigma := sigma + Sqr(A[k,j]/max_abs);
  alpha := Sqrt(sigma);
  sigma := 0.0;
  FOR j := 1 TO r_dim+k DO BEGIN
   IF j = r_dim+k THEN
    IF A[k,j] < 0 THEN
     v[j] := A[k,j]/max_abs - alpha
    ELSE
     v[j] := A[k,j]/max_abs + alpha
   ELSE
    v[j] := A[k,j]/max_abs;
   sigma := sigma + Sqr(v[j]);
  END;
  alpha := 2/sigma;
  FOR i := 1 TO k DO BEGIN
   sigma := 0.0;
   FOR j := 1 TO r_dim+k DO
    sigma := sigma + A[i,j]*v[j];
   beta := alpha * sigma;
   FOR j := 1 TO r_dim+k DO
    A[i,j] := A[i,j] - beta*v[j];
  END;
 END; {for k := n_dim downto 1}
END;
```

APPENDIX 5-2: ALGORITHM FOR COMPUTING THE DERIVATIVE OF A CHOLESKY FACTOR OF A SYMMETRIC MATRIX

```
{Given a matrix dA and the upper triangular Cholesky decomposition CA,
of the original matrix A, d_choldcmp_n calculates the derivative of the
Cholesky decomposition of A. Based on formula dA = dCA.CA' + CA.dCA',
dA,CA known dCA is calculated. Not destructive. Output is in the form of
a pointer to an nxn matrix}
```

```
{NOTES : You have to do your own heap-keeping The pointer result of the
function is initialised but cannot be released/disposed of until the
function has returned. Can be modified to pass parameters by variable
for cheaper stack usage.}

{For CA, dCA lower triangular use d_choldcmp_n
d_choldcmp_n used for solving dF^(-1)=dCF^(-1)'.CF^(-1)+CF^(-1)'.dCF^(-
1) where CF, CF^(-1) is upper triangular
 For CA, dCA upper triangular use d_choldcmp_m
d_choldcmp_m used for solving dPt_p=dCPt_p.CPt_p'+CPt_p.dCPt_p'}

{necessary TYPE definitions :
TYPE   real    = double {optional}
       mxm     = ARRAY[1..m,1..m] OF real;
       nxn     = ARRAY[1..n,1..n] OF real;
       mxmptr  = ^mxm;
       nxnptr  = ^nxn;}

FUNCTION d_choldcmp_n(dA,CA:nxn):nxnptr;

VAR tmp : nxn;
    i,j : byte;
    max_abs : real;

FUNCTION sum(B,C:nxn;i,j,k:byte):real;

VAR l : byte;

BEGIN
 IF k > 0 THEN
  FOR l := 1 TO k DO
   sum := B[i,l]*C[j,l]
 ELSE
   sum := 0.0;
END;


BEGIN
 d_choldcmp_n := New(nxnptr);
 max_abs := 0.0;
 FOR i := 1 TO n DO
  FOR j := 1 TO i DO
   IF abs(CA[i,j]) > max_abs THEN max_abs := abs(CA[i,j]);

 FOR i := 1 TO n DO
  FOR j := 1 TO i DO BEGIN
   dA[i,j] := dA[i,j]/max_abs;     {  re-scale for better stability }
   CA[i,j] := CA[i,j]/max_abs;     {                                }
   IF i=j THEN tmp[i,j]:=(dA[i,j]-(sum(tmp,CA,i,j,j-1)
                          +sum(CA,tmp,i,j,j-1)))/(2*CA[j,j])
    ELSE
   tmp[i,j] := (dA[i,j]-(sum(tmp,CA,i,j,j-
1)+sum(CA,tmp,i,j,j)))/CA[j,j];
  END;
 FOR j := 1 TO n DO
  FOR i := 1 TO j-1 DO
   tmp[i,j] := 0.0;
 d_choldcmp_n^:=tmp;
END;

FUNCTION d_choldcmp_m(dA,CA:mxm):mxmptr;

VAR tmp : mxm;
    i,j : byte;
    max_abs : real;

FUNCTION sum(B,C:mxm;i,j,k:byte):real;

VAR l : byte;

BEGIN
 IF k <= m THEN
```

```
  FOR l := k TO m DO
   sum := B[i,l]*C[j,l]
 ELSE
   sum := 0.0;
END;


BEGIN
 d_choldcmp_m := New(mxmptr);
 max_abs := 0.0;
 FOR i := m DOWNTO 1 DO
  FOR j := m DOWNTO i DO
   IF abs(CA[i,j]) > max_abs THEN max_abs := abs(CA[i,j]);

 FOR i := m DOWNTO 1 DO
  FOR j := m DOWNTO i DO BEGIN
   dA[i,j] := dA[i,j]/max_abs;    {  re-scale for better stability }
   CA[i,j] := CA[i,j]/max_abs;    {                                }
   IF i=j THEN tmp[i,j]:=(dA[i,j]-(sum(tmp,CA,i,j,j+1)
                         +sum(CA,tmp,i,j,j+1)))/(2*CA[j,j])
    ELSE
   tmp[i,j] := (dA[i,j]-
(sum(tmp,CA,i,j,j+1)+sum(CA,tmp,i,j,j)))/CA[j,j];
  END;
 FOR i := 1 TO m DO
  FOR j := 1 TO i-1 DO
   tmp[i,j] := 0.0;
 d_choldcmp_m^:=tmp;
END;
```

6. CONCLUSION

Investigation into the dynamical behaviour of a standard economic model has been able to provide justification for the form of the drift function assumed by, for instance, Hull and White (1994b) and Babbs and Webber (1994). Introduction of a third factor led to a model, the dynamical possibilities of which include chaos. For realistic parameter values, chaotic behaviour is observed. The deterministic driving terms cause switching between high interest rate regimes and low interest rate regimes in a manner suggestive of business cycles. An attempt to reconstruct the attractor shows surprising evidence of structure. A heuristic approach to estimation led to an investigation of a number of methods of exploiting the geometry of the chaotic system and the presence of noise. The use of filtering techniques has been explored. Investigations were conducted into the implications of estimating non-linear models by this method. I was led to propose some new methods for applying filtering to models exhibiting complex dynamics. Estimation results for the three factor model are compromised by the constrictions necessary for the estimation procedure as well as the length of the data series.

The thesis has achieved four objectives. Firstly, it has provided an economic context into which a number of stochastic mean models could be classified. The class of models is quite broad and allows for a variety of dynamical specifications, although it excludes models which incorporate stochastic volatility. The definition of the new class finds most existing models have a trivial form within it. This leaves a wide range of dynamical specifications open to investigation. Secondly, an example of a plausible interest rate model has been motivated. The model is capable of endogenously describing a diversity of behaviour including regime switching and business cycle type behaviour. Thirdly, it has been indicated that methods rooted in the analysis of non-linear systems may be applied directly to term structure data. Fourthly, insights have been gained into how interest rate models exhibiting complex dynamics may be estimated. A variety of heuristic approaches have shown how knowledge of the geometry of a system may provide an estimation approach. The Kalman filter has been shown to be a useful tool in the estimation of complex interest rate models. Advancements over the standard filtering techniques were found to be necessary for dealing with the complex nature of such systems.

6.1 Further Research

This research has sought to investigate several novel aspects of term structure modelling. Little previous work has concentrated on either the economic fundamentals of a system or complex dynamics as deterministic driving components of the large scale dynamics of the short rate process.

From an economic perspective, the economic model in chapter two can be extended significantly. Price dynamics and a supply side, including the labour market, could be introduced. The assumption of a closed economy could be relaxed, and the exchange rate incorporated. The resultant models as expressed in the dynamic mean framework will be more open to economic interpretation and control, as well as better reflecting the real interaction of markets in the economy.

Some discussion is given in the thesis to control of the chaotic model developed in chapter three. There is much scope for investigation of the control of such complex systems. Literature on the control of chaotic systems is well developed. See for instance Abarbanel et al (1993). The implications for control of chaotic systems via control of economic policy variables are intriguing. If authorities are able to control the system, then there is the possibility for them to switch the system from one regime to another, targeting certain rate levels and synchronisation of the interest rate with other fundamentals of the internal or global economy.

The estimation of complex dynamical systems of the short rate has been seen to be aided by the relationship between the geometry of the system and properties of the term structure. Little discussion is given in the thesis to the quantification of complex behaviour in the empirical short rate by test statistics. This is purposefully avoided as they are not the focus of the research and their application to economic/financial series is hampered by the empirical data series available. Typically, calculation of the largest Lyapunov exponent or correlation dimension requires some 10,000 data points. It is also necessary that this data is set over a relatively wide time span, so as to capture the full range of dynamics of the system. Some investigation into the invariants of the dynamics of the empirical process would be beneficial although it is expected that conclusions drawn from such an analysis would be qualified and limited. This assessment is based on the mixed results found in a variety of empirical studies of economic/financial series. See Lorenz (1997) for a survey of such studies.

Techniques introduced in chapter five for advancement of the SRCF algorithm need quantification. Although the revised method for calculating the analytical derivatives of the log-likelihood function is proposed and utilised, the numerical properties of the algorithms involved are not assessed. This is not entirely a trivial problem, and would necessitate an analysis not dissimilar to that by Verhaegen and Van Dooren (1986) in their assessment of the numerical aspects of different Kalman filter implementations. Much of the difficulties observed in the empirical analysis stem from the lack of explicit solutions. Greater understanding of the biases involved in the discretisations employed, as well as approximations to the term structure, would enhance the results. As alluded to in chapter five, the implications for such biases in non-linear (and chaotic) models may be considerably more complex than in the linear case. A further technique of interest in the estimation problem is that of application of non-linear variants of the Kalman filter. In essence, these variants work by computing a linearisation of the system around either the nominal or estimated trajectory. The drawback of these techniques comes in the computational complexity involved. Grewal and Andrews (1993) provide a discussion of these techniques.

Overall, it is thought that the topics motioned above are departures from the central themes of the research. Proper treatment would involve substantive work, and as such, that work is left for future research.

BIBLIOGRAPHY

ABARBANEL HDI, BROWN R, SIDOROWICH JJ AND TSIMRING LS (1993). The analysis of observed chaotic data in physical systems. *Reviews of Modern Physics* **65**: 1331-1392.

ABHYANKAR A, COPELAND LS AND WONG W (1997). Uncovering nonlinear structure in real-time stock market indices: the S&P 500, the DAX, the Nikkei 225 and the FTSE-100. *Journal of Business and Economic Statistics* **15**:1-14.

AHN CM AND THOMPSON HE (1988). Jump diffusion processes and the term structure of interest rates. *The Journal of Finance* **48**: 155-174.

AÏT-SAHALIA Y (1996). Testing continuous-time models of the spot interest rate. *Review of Financial Studies* **9**: 385-426.

ALEXANDER C AND GIBLIN I (1994). Chaos in the system. *Risk Magazine* **7**.

AMMERAAL L (1995). *C++ for Programmers, 2nd ed.* Wiley: Chichester.

ANDERSEN TG AND LUND J (1997). Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics* **77**: 343-377.

ANSELIN L (1988). *Spatial Econometrics: Methods and Models.* Kluwer: Dordrecht.

APARICIO ACOSTA FM (1996). On some chaos techniques and the modelling of nonlinear time series. *Signal Processing* **55**: 269-283.

APOSTOL TM (1974). *Mathematical Analysis, 2nd ed.* Addison-Wesley: Reading.

ARNOLD L (1996). Trends and open problems in the theory of random dynamical systems. Research Paper 365, Institut für Dynamische Systeme, Universität Bremen.

ARROWSMITH DK AND PLACE CM (1995). *Dynamical Systems.* Chapman and Hall: London.

BABBS S (1993). 'Generalised Vasicek' models of the term structure. In: JANSSEN J AND SKIADAS CH (EDS). *Proceedings of the Sixth Annual Symposium on Applied Stochastic Models and Data Analysis* **1**: 49-62. World Scientific: Singapore.

BABBS S AND WEBBER N (1994). A theory of the term structure with an official short rate. Research Paper, Financial Options Research Centre, University of Warwick.

BABBS S AND WEBBER NJ (1997). Term structure modelling under alternative official regimes. In: DEMPSTER MAH AND PLISKA SR (EDS). *Mathematics of Derivative Securities*. Cambridge University Press: Cambridge.

BABBS SH AND NOWMAN KB (1997). Kalman filtering of generalised Vasicek term structure models. Available from: Dr. K. B. Nowman, First National Bank of Chicago, First Chicago House, 90 Long Acre, London WC2E 9RB, UK.

BAEK EG AND BROCK WA (1992). A nonparametric test for independence of a multivariate time series. *Statistics Sinica* **2**: 137-156.

BAKER GL AND GOLLUB JP (1996). *Chaotic Dynamics: an Introduction, 2nd ed.* Cambridge University Press: Cambridge.

BAKSHI GS AND CHEN Z (1997). An alternative valuation equation for contingent claims. *Journal of Financial Economics* **44**: 123:165.

BALDUZZI P, BERTOLA G AND FORESI S (1993). A model of target changes and the term structure of interest rates. National Bureau of Economic Research, working paper 4347

BALL CA AND TOROUS WN (1996) Unit roots and the estimation of interest rate dynamics. *Journal of Empirical Finance*. **3**: 215-238

BARNDORFF-NIELSEN OE, JENSEN JL AND KENDALL WS (EDS) (1993). *Networks and Chaos-Statistical and Probabilistic Aspects*. Chapman & Hall: London.

BARNETT WA (1993). Has chaos been discovered with economic data? In: DAY R AND CHEN P (EDS). *Nonlinear Dynamics and Evolutionary Economics.* Oxford University Press: Oxford.

BARNETT WA, GALLANT AR, HINICH MJ, JUNGEILGES JA, KAPLAN DT AND JENSEN MJ (1995). Robustness of nonlinearity and chaos tests to measurement error, inference method and sample size. *Journal of Economic Behaviour and Organisation* **27**: 301-320.

BARRETT MD (1996). Continuous control of chaos. *Physica D* **91**: 340-348.

BEAGLEHOLE D AND TENNEY M (1991). General solutions of some interest rate contingent claim pricing equations. *Journal of Fixed Income* **1**: 69-83.

BECKER KH AND DORFLER M (1989). *Dynamical Systems and Fractals.* Cambridge University Press: Cambridge.

BERGÉ P AND DUBOIS M (1983). Transient re-emergent order in convective spatial chaos. *Physics Letters A* **93**: 365-367.

BERGSTROM AR (1990). *Continuous Time Econometric Modelling.* Oxford University Press: Oxford.

BIERMAN GJ AND THORNTON CL (1977) Numerical comparisons of Kalman filter algorithms: Orbit determination case study. *Automatica* **13**: 23-35

BISHOP AR, FESSER K, LOMDAHL PS, KERR WC, WILLIAMS MB AND TRULLINGER SE (1983). Coherent spatial structure versus time chaos in a perturbed Sine-Gordon system. *Physical Review Letters* **50**: 1095-1098.

BJÖRK T (1996). Interest Rate Theory. CIME Lectures. To appear in *Springer Lecture Notes in Mathematics.*

BLACK D AND DOWD M (1994). The money multiplier, the money market and the LM curve. *Eastern Economic Journal* **20**: 301-310.

BLACK F AND KARASINSKI P (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal* **47**: 52-59.

BLACK F, DERMAN E AND TOY W (1990). A one factor model of interest rates and its application to treasury bond options. *Financial Analysts Journal* **Jan-Feb**: 33-39.

BLACK F, SCHOLES M (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* **81**: 637-54.

BLANCHARD OJ AND FISCHER S (1990). *Lectures on Macroeconomics.* Cambridge University Press: Cambridge.

BLANK SC (1991). 'Chaos' in futures markets? A nonlinear dynamical analysis. *The Journal of Futures Markets* **11**: 711-728.

BLENMAN LP, CANTRELL RS, FENNELL RE, PARKER DF, RENEKE JA, WANG LFS AND WOMER NK (1995). An alternative approach to stochastic calculus for economic and financial models. *Journal of Economic Dynamics and Control* **19**: 553-568.

BODIE Z, KANE A AND MARCUS AJ (1989). *Investments.* Irwin: Homewood.

BOLLERSLEV T, CHOU RY AND KRONER KF (1992). ARCH modelling in finance: a review of theory and empirical evidence. *Journal of Econometrics* **52**: 5-59.

BOROWSKI EJ AND BORWEIN JM (1989). *Dictionary of Mathematics.* Harper Collins.

BOX GEP JENKINS GM AND REINSEL GC (1994). Time Series Analysis: Forecasting and Control, 3rd Ed. Prentice Hall: Englewood Cliffs.

BRANSON WH (1989). *Macroeconomic Theory and Policy, 3rd ed.* Harper & Row: New York.

BRENNAN MJ AND SCHWARTZ ES (1979). A continuous time approach to the pricing of bonds. *Journal of Banking and Finance* **3**: 133–155.

BRENNAN MJ AND SCHWARTZ ES (1980). Analysing convertible bonds. *Journal of Financial and Quantitative Analysis* **15**: 907-929.

BRENNAN MJ AND SCHWARTZ ES (1982). An equilibrium model of bond pricing and a test of market efficiency. *Journal of Financial and Quantitative Analysis* **18**: 301-329.

BROCK WA (1986). Distinguishing random and deterministic systems: abridged version. In: GRANDMONT J-M (ED). *Nonlinear Economic Dynamics*: 168-195. Academic Press: Orlando.

BROCK WA (1986). Is the business cycle characterised by deterministic chaos? In: GRANDMONT J (ED). *Nonlinear Economic Dynamics.* Academic Press: Orlando.

BROCK WA (1988). Nonlinearity and complex dynamics in economics and finance. In: ANDERSON PW, ARROW KJ AND PINES D (EDS). *The Economy as an Evolving Complex System.* Addison Wesley: Reading.

BROCK WA (1993). Pathways to randomness in the economy: emergent nonlinearity and chaos in economics and finance. *Estudios Económicos* **8**: 3-56.

BROCK WA (1996). Asset Price Behaviour in Complex Environments. Research Paper, Social Systems Research Institute, University of Wisconsin Madison.

BROCK WA AND BAEK EG (1991). Some theory of statistical inference for nonlinear science. *Review of Economic Studies* **58**: 697-716.

BROCK WA AND BAEK EG (1992). A nonparametric test for independence of a multivariate time series. *Statistica Sinica* **2**: 137-156.

BROCK WA AND DE LIMA PJF (1996). Nonlinear time series, complexity theory, and finance. In: MADDALA GS AND RAO CR (EDS). *Handbook of Statistics* **14**: 317-361. Elsevier Science B. V.

BROCK WA AND HOMMES CH (1995). Rational routes to randomness. Research Paper, Social Systems Research Institute, University of Wisconsin-Madison.

BROCK WA AND HOMMES CH (1997). Models of complexity in economics and finance. Research Paper, Social Systems Research Institute, University of Wisconsin Madison.

BROCK WA AND KLEIDON AW (1992). Periodic market closure and trading volume - a model of intraday bids and asks. *Journal of Economic Dynamics and Control* **16**: 451.

BROCK WA AND LEBARON BD (1996). A dynamic structural model for stock return volatility and trading volume. *Review of Economics and Statistics* **78**: 94-110.

BROCK WA AND MALLIARIS AG (1989). *Differential Equations, Stability and Chaos in Dynamic Economics*. North Holland: Oxford.

BROCK WA AND POTTER SM (1993). Nonlinear time series and macroeconometrics. In: RAO CR (ED). *Handbook of Statistics* **11**. Elsevier Science.

BROCK WA AND SAYERS CL (1988). Is the business cycle characterized by deterministic chaos? *Journal of Monetary Economics* **22**: 71-90.

BROCK WA, DECHERT WD, LEBARON B AND SCHEINKMAN JA (1996). A test for independence based on the correlation dimension. *Econometric Reviews* **15**: 197-235.

BROCK WA, HSIEH DA AND LEBARON B (1991). *Nonlinear Dynamics, Chaos and Instability: Statistical Theory and Economic Evidence*. MIT Press: London.

BROCK WA, LAKONISHOK J AND LEBARON B (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance* **47**.

BROOMHEAD DS AND KING GP (1986). Extracting qualitative dynamics from experimental data. *Physica D* **20**: 217-236.

BROWN R AND CHUA LO (1996). Clarifying chaos: examples and counterexamples. *International Journal of Bifurcation and Chaos* **6**: 219:249.

BROWN R, RULKOV N AND TUFILLARO N (1994). The effects of additive noise and drift in the dynamics of the driving on chaotic synchronisation. Unpublished Paper, Institute for Nonlinear Science, University of California, San Diego and Centre for Nonlinear Studies, Los Alamos National Laboratory.

BROWN RH AND SCHAEFER SM (1994). The term structure of real interest rates and the Cox, Ingersoll and Ross model. *Journal of Financial Economics* **35**: 3-42.

BUCY RS AND JOSEPH PD (1968) *Filtering for Stochastic Processes, with Applications to Guidance*. Wiley, New York, CAGINALP G AND BALENOVICH D (1994). Market oscillations induced by the competition between value-based and trend-based investment strategies. *Applied Mathematical Finance* **1**: 129-164.

CANABARRO E (1993). Comparing the dynamic accuracy of yield-curve-based interest rate contingent claim pricing models. *Journal of Financial Engineering* **2**: 365-401.

CAO CQ AND TSAY RS (1992). Non-linear time series analysis of stock volatilities. *Journal of Applied Econometrics* **7**: 165-185.

CARROLL RJ, RUPPERT D AND STEFANSKI LA (1995). *Measurement Error in Nonlinear Models.* Chapman & Hall: London.

CHAPMAN DA, PEARSON ND (1998). Is the short rate drift actually nonlinear? Working Paper, Graduate School of Business, University of Texas. Available from http://www.bus.utexas.edu/~chapmand/cp.html.

CHAPMAN DA, LONG JB AND PEARSON ND (1998). Using proxies for the short rate: When are three months like an instant? Working paper, Finance Department, Graduate School of Business, University of Texas. Available from http://www.bus.utexas.edu/~chapmand/clp.html.

CHAN KC, KAROLYI C, LONGSTAFF GA AND SANDERS FA (1992). The volatility of short term interest rates: an empirical comparison of alternative models of the term structure of interest rates. *Journal of Finance* **47**: 209-1227.

CHEN L (1996). Stochastic mean and stochastic volatility - a three-factor model of the term structure of interest rates and its implications in derivatives and pricing management. *Financial Markets, Institutions and Instruments* **5**: 1-88.

CHEN P (1993). Searching for economic chaos: a challenge to econometric practice and nonlinear tests. In: DAY R AND CHEN P (EDS). *Nonlinear Dynamics and Evolutionary Economics.* Oxford University Press: Oxford.

CHEN S AND BILLINGS A (1992). Neural networks for nonlinear dynamic systems modelling and identification. *International Journal of Control* **56**: 319-346.

CHIANG AC (1984). *Fundamental Methods of Mathematical Economics, 3rd Ed.* McGraw-Hill: Singapore.

CHICHILNISKY G, HEAL G AND LIN Y (1995). Chaotic price dynamics, increasing returns and the Phillips curve. *Journal of Economic Behaviour and Organisation* **27**: 279-291.

CHOI H (1995). Goodwin's growth cycle and the efficiency wage hypothesis. *Journal of Economic Behaviour and Organisation* **27**: 223-35.

COPELAND TE AND WESTON JP (1988). *Financial Theory and Corporate Policy.* Addison-Wesley: Reading.

COULLET P, ELPHICK C AND REPAUX D (1987). Nature of spatial chaos. *Physical Review Letters* **58**: 431-435.

COX DR AND MILLER HD (1995). *The Theory of Stochastic Processes.* Chapman & Hall: London.

COX JC, INGERSOLL JE AND ROSS SA (1980). An analysis of variable rate loans contracts. *The Journal of Finance* **35**: 389-403.

COX JC, INGERSOLL JE AND ROSS SA (1985a). An intertemporal general equilibrium model of asset prices. *Econometrica* **53**: 363-384.

COX JC, INGERSOLL JE AND ROSS SA (1985b). A theory of the term structure of interest rates. *Econometrica* **53**: 385-407.

CREEDY J AND MARTIN VL (EDS) (1994). *Chaos and Non-linear Models in Economics.* Edward Elgar: Aldershot.

DAY RH (1992). Complex economic dynamics: obvious in history, generic in theory, elusive in data. *Journal of Applied Econometrics* **7**: S9-S23.

DAY RH (1994). *Complex Economic Dynamics.* MIT Press: Cambridge.

DAY RH AND CHEN P (1993). *Nonlinear Dynamics and Evolutionary Economics.* Oxford University Press: Oxford.

DAY RH AND SHAFER W (1992). Keynesian chaos. In: BENHABIB J (ED). *Cycles and Chaos in Equilibrium.* Princeton University Press: Princeton.

DE GRAWE P, DEWACHTER H AND EMBRECHTS M (1994). *Exchange Rate Theory: Chaotic Models of Foreign Exchange Markets.* Blackwell: Oxford.

DE JONG P AND SHEPARD N (1995). The simulation smoother for time series models. *Biometrika* **82**: 339-350.

DECHERT WD AND GENCAY R (1992). Lyapunov exponents as a nonparametric diagnostic for stability analysis. *Journal of Applied Econometrics* **7**: S41-S60.

DEISSLER RJ (1987). Spatially growing waves, intermittency, and convective chaos in an open-flow system. *Physica D* **25**: 233-260.

DERNBURG TF AND DERNBURG JD (1969). *Macroeconomic Analysis: An Introduction to Comparative Statics and Dynamics.* Addison-Wesley: Reading.

DEVANEY RL (1989). *An Introduction to Chaotic Dynamical Systems, 2nd Ed.* Addison-Wesley: Reading.

DEWSON T AND IRVING AD (1996). Mixed order response function estimation from multi-input non-linear systems. *Physica D* **94**: 19-35.

DORNBUSCH R AND FISCHER S (1990). *Macroeconomics, 5th Ed.* McGraw-Hill: Singapore.

DOTHAN LU (1978). On the term structure of interest rates. *Journal of Financial Economics* **6**: 59-69.

DOTHAN MU (1990). *Prices in Financial Markets.* Oxford University Press: Oxford.

DUFFIE D (1992). *Dynamic Asset Pricing Theory.* Princeton University Press: Princeton.

DUFFIE D AND KAN R (1994). Multi-factor term structure models. *Philosophical Transactions of the Royal Society of London A* **347**: 577-586.

DUFFIE D AND KAN R (1996). A yield factor model of interest rates. *Mathematical Finance* **6**: 379-406.

ECKMANN J-P AND RUELLE D (1985). Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics* **57**: 617-656.

EDELHART M AND NATH S (1988). *The Brady Book of Turbo Pascal.* Prentice Hall: New York.

EDGAR GA (1992). *Measure, Topology, and Fractal Geometry.* Springer-Verlag: New York.

ELTON EJ AND GRUBER MJ (1991). *Modern Portfolio Theory and Investment Analysis.* Wiley: New York.

ENGLE RF (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**: 987-1007.

EVANS GA (1995). *Practical Numerical Analysis.* Wiley: Chichester.

FARMER REA (1986). Deficits and cycles. In: GRANDMONT J-M (ED). *Nonlinear Economic Dynamics*: 77-88. Academic Press: Orlando.

FRANKEL JA (1993). *On Exchange Rates.* MIT Press: Cambridge.

FRAZER AM (1996). Chaos and detection. *Physical Review E* **53**: 4514: 4523.

FROYLAND G, JUDD K AND MEES AI (1995). Estimation of Lyapunov exponents of dynamical systems using a spatial average. *Physical Review E* **51**: 2844-2855.

GENÇAY R (1996). A statistical framework for testing chaotic dynamics via Lyapunov exponents. *Physica D* **89**: 261-266.

GILMORE CG (1993). A new test for chaos. *Journal of Economic Behaviour and Organization* **22**: 209-237.

GOLUB G E AND VAN LOAN CF (1983). *Matrix Computations.* North Oxford Academic: Oxford.

GÓMEZ V (undated). Estimation and signal extraction for finite nonstationary series with the Kalman filter. Research Paper, Available from: Víctor Gómez, Ministerio de Economía y Hacienda, Paseo de la Castellana 162, 28046 Madrid, Spain.

GOODWIN RM (1990). *Chaotic Economic Regimes.* Clarendon: Oxford.

GOODWIN RM AND PACINI PM (1992). Nonlinear economic dynamics and chaos: an introduction. In: VERCELLI A AND DIMITRI N (EDS). *Macroeconomics: A Survey of Research Strategies*: 236-291. Oxford University Press: Oxford.

GOURIÉROUX C AND MONFORT A (1996). *Simulation-Based Econometric Methods.* Oxford University Press: Oxford.

GRANDMONT J-M (1986). Stabilising competitive business cycles. In: GRANDMONT J-M (ED). *Nonlinear Economic Dynamics*: 57-76. Academic Press: Orlando.

GRANGER CW AND TERÄSVIRTA T (1993). *Modelling Nonlinear Economic Relationships.* Oxford University Press: Oxford.

GRANGER CWJ (1968). Some aspects of the Random Walk Model of stock market prices. *International Economic Review* **9**: 253-259.

GRANGER CWJ (1992). Forecasting stock market prices: lessons for forecasters. *International Journal of Forecasting* **8**: 3-13.

GRANGER CWJ (ED) (1990). *Modelling Economic Series, Readings in Econometric Methodology.* Clarendon: Oxford.

GRANGER CWJ (ED) (1991). *Long Run Economic Relationships, Readings in Cointergration.* Oxford University Press: Oxford.

GRANGER CWJ AND NEWBOLD P (1986). *Forecasting Economic Time Series, 2nd Ed.* Academic Press: Orlando.

GRASSBERGER P AND PROCACCIA I (1983). Measuring the strangeness of the strange attractor. *Physica D* **9**: 189-208.

GRENFELL BT, KLECZKOWSKI A, ELLNER SP AND BOLKER BM (1994). Measles as a case study in nonlinear forecasting and chaos. *Philosophical Transactions of the Royal Society of London A* **348**: 515-530.

GREWAL MS AND ANDREWS AP (1993). *Kalman Filtering: Theory and Practice.* Prentice Hall: Englewood Cliffs.

GU M (1993). An empirical examination of the deterministic component in stock price volatility. *Journal of Economic Behaviour and Organisation* **22**: 239-252.

GU M (1995). Market mediating behaviour: an economic analysis of security exchange specialists. *Journal of Economic Behaviour and Organisation* **27**: 237-56.

GUILLAUME DM, PICTET OV, MÜLLER UA AND DACOROGNA MM (1995). Unveiling non-linearities through time scale transformations. Research Paper, Olsen Associates, Available from http:\\www.olsen.ch\

GUJARATI DN (1988). *Basic Econometrics, 2nd Ed.* McGraw-Hill: Singapore.

HAKKIO CS AND LEIDERMAN L (1986). Intertemporal asset pricing and the term structure of exchange rates and interest rates - the eurocurrency market. *European Economic Review* **30**: 325-344.

HAMILTON JD (1994). *Time Series Analysis.* Princeton University Press: Princeton.

HARVEY AC (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press: Cambridge.

HARVEY AC AND STOCK JH (1993). Estimation, smoothing, interpolation, and distribution for structural time-series models in continuous time. In: PHILLIPS PCB (ED). *Modern Methods and Applications of Econometrics.* Blackwell: Oxford.

HE X AND LAPEDES A (1994). Nonlinear modelling and prediction by successive approximation using radial basis functions. *Physica D* **70**: 289-301.

HEATH D, JARROW R AND MORTON A (1990). Bond pricing and the term structure of interest rates: a discrete time approximation. *Journal of Financial and Quantitative Analysis* **25**: 419-440.

HEATH D, JARROW R AND MORTON A (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica* **60**: 77-105.

HICKS JR (1937). Mr. Keynes and the 'Classics': a suggested interpretation. *Econometrica* **5**: 147-159.

HILBORN RC (1994). *Chaos and Non-linear Dynamics.* Open University Press: Oxford.

HOEL PG (1984). *Introduction to Mathematical Statistics, 5th Ed.* Wiley: New York.

HOGAN M. (1993). Problems in certain two-factor term structure models. *Annals of Applied Probability* **3**: 576-581.

HOLT RG AND HOLT DL (1994). On defining chaos. Research Paper, Yale University and University of the South.

HOLYST JA, TAGEL T, HAAG G AND WEIDLICH W (1996). How to control a chaotic economy? *Journal of Evolutionary Economics* **6**: 31-42.

HONORÉ P (1998). Five essays on financial econometrics in continuous-time models. PhD Thesis, The Aarhus School of Business, Denmark.

HSIEH DA (1991). Chaos and nonlinear dynamics - applications to financial markets. *Journal of Finance* **46**: 1839-1877.

HSEIH SF, LIU KJR AND YAO K (1993) A unified square-root-free approach for QRD-based recursive least squares estimation. *IEEE Transactions on Signal Processing.* **41(3)**: 1405-1409

HUANG W (1995). Caution implies profit. *Journal of Economic Behaviour and Organisation* **27**: 257-277.

HULL J AND WHITE A (1990a). Pricing interest rate derivative securities. *Review of Financial Studies* **3**: 573-592.

HULL J AND WHITE A (1990b). Valuing derivative securities using the explicit finite difference method. *Journal of Financial and Quantitative Analysis* **25**: 87-100.

HULL J AND WHITE A (1993). One factor interest rate models and the valuation of interest rate derivative securities. *Journal of Financial and Quantitative Analysis* **28**: 235-254.

HULL J AND WHITE A (1993). The pricing of options on interest-rate caps and floors using the Hull-White Model. *Journal of Financial Engineering* **2**: 287-296.

HULL J AND WHITE A (1994a). Numerical procedures for implementing Term Structure Models I: single-factor models. *Journal of Derivatives* **Fall**: 7-16.

HULL J AND WHITE A (1994b). Numerical procedures for implementing Term Structure Models II: two-factor models. *Journal of Derivatives* **Winter**: 37-48.

HULL JC (1993). *Options, Futures and Other Derivative Securities, 2nd Ed.* Prentice-Hall: London.

INGERSOLL JE (1987). An introduction to stochastic calculus. In: INGERSOLL JE. *Theory of Financial Decision Making*: 347-360. Rowman & Littlefield: Totowa.

IRVING AD AND DEWSON T (1996). Generalised autoregression and the analysis of dynamical processes. *Applied Mathematical Modelling* **20**: 152-161.

IRVING AD, CUNLIFFE NH, DEWSON T AND HONG G (1995). General nonlinear response of a single input system to stochastic excitations. *Applied Mathematical Modelling* **19**: 46-55.

JAMSHIDIAN F (1991). Forward induction and construction of yield curve models. Financial Strategies Group, Merill Lynch.

JAMSHIDIAN F (1993). Bond, futures and option valuation in the quadratic interest rate model. Research Paper, Fuji International Finance.

JEGADEESH, N. AND PENNACCHI GG (1996) The behaviour of interest rates implied by the term structure of Eurodollar Futures. *Journal of Money Credit and Banking.* **28:** 426-446.

JOHNSTON J (1984). *Econometric Methods, 3rd Ed.* McGraw-Hill: Singapore.

JUDD K (1994). Estimating dimension from small samples. *Physica D* **71**: 421-429.

JUDGE GE, GRIFFITHS WE, CARTER HILL R, LÜTKEPOHL H AND LEE T-C (1985). *The Theory and Practice of Econometrics, 2nd Ed.* Wiley: New York.

KANEKO K (1989). Pattern dynamics in spatiotemporal chaos. *Physica D* **34**: 1-41.

KARATZAS I AND SHREVE SE (1991). *Brownian Motion and Stochastic Calculus, 2nd Ed.* Springer-Verlag: New York.

KENNEDY DP (1994) Characterising and filtering gaussian models of the term structure of interest rates. Working paper, Statistical Laboratory, University of Cambridge

KITAGAWA G (1981). A nonstationary time series model and its fitting by a recursive filter. *Journal of Time Series Analysis* **2**: 103-116.

KITAGAWA G AND HIROTUGU A (1978). A procedure for the modelling of non-stationary time series. *Annnals of Institute of Statistical Mathematics B* **30**: 351-363.

KOSTELICH EJ (1992). Problems in estimating dynamics from data. *Physica D* **58**: 138-152.

KRUGMAN P AND MILLER M (EDS) (1994). *Exchange Rate Targets and Currency Bands.* Cambridge University Press: Cambridge.

KWASNIOK F (1996). The reduction of complex dynamical systems using principal interaction patterns. *Physica D* **92**: 28-60.

LAMBERT PJ (1992). *Advanced Mathematics for Economists.* Blackwell: Oxford.

LANGETIEG, T. (1980). A multivariate model of the term structure. *Journal of Finance.* **35**: 71-97.

LARRAIN M (1991). Testing chaos and non-linearities in T-Bill rates. *Financial Analysts Journal* **Sep-Oct**: 51-62.

LARRAIN M AND PAGANO M (1993). Forecasts from a non-linear T-Bill rate model. *Financial Analysts Journal* **Nov-Dec**: 83-88.

LEBARON B (1992). Forecast improvements using a volatility index. *Journal of Applied Econometrics* **7**: S137-S149.

LEBARON B (1994). Chaos and nonlinear forecastability in economics and finance. *Philosophical Transactions of the Royal Society of London A* **348**: 397-404.

LEBARON B (1994). Technical trading rule profitability and foreign exchange intervention. Unlisted Research Paper, Department of Economics, University of Wisconsin-Madison.

LEBARON B AND WEIGEND AS (1994). Evaluating neural network predictors by bootsrapping. Unlisted Research Paper, Department of Economics, University of Wisconsin-Madison, and Department of Computer Science and Institute of Cognitive Science, University of Colorado.

LEVY A (1992). *Economic Dynamics.* Avebury: Aldershot.

LINDEN N, SATCHELL S AND YOON Y (1993). Predicting British financial indices: an approach based on Chaos Theory. *Structural Change and Economic Dynamics* **4**: 145-162.

LIU T, GRANGER CWJ AND HELLER WP (1992). Using the correlation exponent to decide whether an economic series is chaotic. *Journal of Applied Econometrics* **7**: S25-S39.

LONGSTAFF FA (1989). A nonlinear general equilibrium model of the term structure of interest rates. *Journal of Financial Economics* **23**: 195-224.

LONGSTAFF FA (1992). Multiple equilibria and term structure models. *Journal of Financial Economics* **32**: 333-344.

LONGSTAFF FA AND SCHWARTZ ES (1992). Interest rate volatility and the term structure: a two factor general equilibrium model. *The Journal of Finance* **47**: 1259-1282.

LORENZ H-W (1997). *Nonlinear Dynamical Economics and Chaotic Motion, 2nd Ed.* Springer-Verlag: Berlin.

LORETO V, PALADIN G AND VULPIANI A (1996). Concept of complexity in random dynamical systems. *Physical Review E* **53**: 2087-2098.

LUND J. (1997) Econometric analysis of continuous time arbitrage-free models of the term structure of interest rates. Working paper, Department of Finance, Aarhus School of Business, Denmark.

MACDONALD R AND TAYLOR MP (1992). Exchange rate economics - a survey. *IMF Staff Papers* **39**: 1-57.

MACKAY RS (1992). Nonlinear dynamics in economics: a review of some key features of nonlinear dynamical systems. Research Paper, Nonlinear Systems Laboratory, Mathematics Institute, University of Warwick.

MACKAY RS (1993). On transport in Hamiltonian systems. Research Paper, Nonlinear Systems Laboratory, Mathematics Institute, University of Warwick.

MACKAY RS (undated). Some thoughts on chaos in engineering. Unpublished Research Paper, Nonlinear Systems Laboratory, Mathematics Institute, University of Warwick.

MALLIARIS AG AND WA BROCK (1982). *Stochastic Methods in Economics and Finance*. North-Holland: Amsterdam.

MAREK M AND SCHREIBER I (1991). *Chaotic Behaviour of Deterministic Dissipative Systems*. Cambridge University Press: Cambridge.

MARTINERIE JM, ALBANO AM, MEES AI AND RAPP PE (1992). Mutual information, strange attractors, and the optimal estimation of dimension. *Physical Review A* **45**: 7058-7064.

MASTERS T (1995). *Neural, Novel and Hybrid Algorithms for Time Series Production.* Wiley: New York.

MASTRONIKOLA (1991). Yield curve for gilt edge stocks: A new model. Working Paper, Bank of England

MAYFIELD ES AND MIZRACH B (1991). Nonparametric estimation of the correlation exponent. *Physical Review A* **44**: 5298-5301.

MCNEVIN B AND NEFTÇI S (1992). Some evidence of non-linearity of economic time series: 1890-1981. In: BENHABIB J (ED). *Cycles and Chaos in Economic Equilibrium*: 429-445. Princeton University Press: Princeton.

MEDIO A AND GALLO G (1993). *Chaotic Dynamics, Theory and Applications to Economics.* Cambridge University Press: Cambridge.

MERTON RC (1992). *Continuous-Time Finance, Rev. Ed.* Blackwell: Oxford.

MIZRACH B (1992). Multivariate nearest neighbour forecasts of EMS exchange rates. *Journal of Applied Econometrics* **7**: S151-S163.

MIZRACH B (1992). The state of economic dynamics: a review essay. *Journal of Economic Dynamics and Control* **16**: 175-190.

MIZRACH B (1995). Using U-statistics to detect business cycle nonlinearities. In: SEMMLER W (ED). *Business Cycles: Theory and Empirical Methods.* Kluwer: London.

MIZRACH B (1996). The information in the term structure: a non-parametric investigation. *Journal of Forecasting* **15**: 137-154.

MORF M AND KAILATH T (1975) Square root algorithms for least squares estimation. *IEEE Transactions on Automatic Control.* **20:** 487-497

MULDOON MR, MACKAY RS, HUKE JP AND BROOMHEAD DS (1993). Topology from time series. *Physica D* **65**: 1-16.

MÜLLER UA, DACOROGNA MM, DAVÉ RD, PICTET OV, OLSEN RB AND WARD JR (1995). Fractals and intrinsic time - a challenge to econometricians. Research Paper, Olsen Associates.

MULLIN T (ED) (1995). *The Nature of Chaos.* Oxford University Press: Oxford.

MULLINEAUX A AND PENG W (1993). Nonlinear business cycle modelling. *Journal of Economic Surveys* **7**: 41-83.

NASH JC (1979). *Compact Numerical Methods for Computers.* Adam Hilger: Bristol.

NIELSEN LT AND SAÁ-REQUEJO J (1992). Exchange rate and term structure dynamics and the pricing of derivative securities. Research Paper, INSEAD, France.

NISHIMURA K AND YANO M (1995). Durable capital and chaos in competitive business cycles. *Journal of Economic Behaviour and Organisation* **27**: 165-81.

NUNES J AND WEBBER N (1996). Low dimensional dynamics and the stability of HJM term structure models. Research Paper, Available from: Nick Webber, Warwick Business School, University of Warwick, Coventry CV4 7AL, UK.

ØKSENDAL B (1992). *Stochastic Differential Equations, 3rd Ed.* Springer-Verlag: Berlin.

PARK P AND KAILATH T (1995). New square-root algorithms for Kalman filtering. *IEEE Transactions on Automatic Control.* **40(5)**: 895-899

PENNACCHI GG (1991) Identifying the dynamics of real interest rates and inflation: Evidence using survey data. *Review of Financial Studies.* **4**: 53-86

PENTECOST EJ (1993). *Exchange Rate Dynamics.* Edward Elgar: Aldershot.

PERKO L (1993). *Differential Equations and Dynamical Systems.* Springer-Verlag: New York.

PESARAN MH AND POTTER SM (1992). Nonlinear dynamics and econometrics: an introduction. *Journal of Applied Econometrics* **7**: S1-S7.

PESARAN MH AND TIMMERMANN A (1994). The use of recursive model selection strategies in forecasting stock returns. Research Paper, Birkbeck College, London.

PETERS EE (1991). A chaotic attractor for the S&P 500. *Financial Analysts Journal* **Mar-Apr**: 55-62.

PETERS EE (1994). *Fractal Market Analysis.* Wiley: New York.

PFANN GA, SCHOTMAN PC AND TSCHERNIG R (1996). Nonlinear interest rate dynamics and implications for the term structure. Journal of Econometrics **74**: 149-176.

PICTET OV, DACOROGNA MM, CHOPARD B, OUSSAIDENE M, SCHIRRU R AND TOMASSINI M (1995). Using genetic algorithms for robust optimization in financial applications. Research Paper, Olsen Associates.

PLATTEN I (1994). Non-linear general equilibrium models of the term structure: comments and two-factor generalizations. *Finance* **15**: 63-78.

PORTER DP AND SMITH VL (1994). Stock market bubbles in the laboratory. *Applied Mathematical Finance* **1**: 111-127.

PRESS WH, FLANNERY BP, TEUKOLSKY SA AND VETTERLING WT (1989). *Numerical Recipes in Pascal.* Cambridge University Press: Cambridge.

PRESS WH, TEUKOLSKY SA, VETTERLING WT AND FLANNERY BP (1992). *Numerical Recipes in C, 2nd Ed.* Cambridge University Press: Cambridge.

RAMSEY JB, SAYERS CL, ROTHMAN P (1992). The statistical properties of dimension calculations using small data sets: some economic applications. In: BENHABIB J (ED). *Cycles and Chaos in Economic Equilibrium*: 394-428. Princeton University Press: Princeton.

RAND DA (1994). Measuring and characterizing spatial patterns, dynamics and chaos in spatially extended dynamical systems and ecologies. *Philosophical Transactions of the Royal Society of London A* **348**: 497-514.

RIPLEY BD (1977). Modelling spatial patterns. *Journal of the Royal Statistical Society B* **39**: 172-212.

ROGERS LCG (1997). The potential approach to the term structure of interest rates and foreign exchange rates. *Mathematical Finance* **7**: 157-176.

ROMA A AND TOROUS W (1997). The cyclical behavior of interest rates. *Journal of Finance* **52**: 1519-1542.

ROSS SA (1987). The interrelations of finance and economics: theoretical perspectives. *AEA Papers and Proceedings* **77**: 29-41.

ROTHENBERG TJ (1973). *Efficient Estimation with A Priori Information.* Yale University Press: New Haven.

ROWLANDS G AND SPROTT JC (1992). Extraction of dynamical equations from chaotic data. *Physica D* **58**: 251-259.

SASAKURA K (1995). Political economic chaos. *Journal of Economic Behaviour and Organisation* **27**: 213-221.

SATCHELL S AND A TIMMERMAN (1995). An assessment of the economic value of non-linear foreign exchange rate forecasts. *Journal of Forecasting* **14**: 477-498.

SAVIT R (1988). When random is not random: an introduction to chaos in market prices. *The Journal of Futures Markets* **8**: 271-289.

SAYERS CL (1995). Testing for chaos and non-linearities in macroeconomic time series. In: SEMMLER W (ED). *Business cycles: Theory and empirical methods.* Kluwer: London.

SCHAEFER S (1981) Measuring a tax specific term structure of interest rates in the market for British government securities. *The Economic Journal.* **91**: 415-438

SCHEINKMAN JA AND LEBARON B (1992). Nonlinear dynamics and stock returns. In: BENHABIB J (ED). *Cycles and Chaos in Equilibrium*: 446-474. Princeton University Press: Princeton.

SCHEINKMAN JA AND LEBARON B (1992). Nonlinear dynamics and stock returns. *Journal of Business* **62**: 311-338.

SCHLÖGL E AND SOMMER D (1997). Factor models and the shape of the term structure. Research Paper, Department of Statistics, University of Bonn.

SCHMALFUSS B (1996). The random attractor of the stochastic Lorenz system. Research Paper 367, Institut für Dynamische Systeme, Universität Bremen.

SCHMIDT S.F. (1966), Applications of State Space Methods to Navigation Problems. *Advances in Control Systems,* **3**, Academic Press, San Diego

SCHMIDT S.F. (1981) The Kalman filter: Its recognition and development for aerospace applications. *AIAA Journal of Guidance and Control.* **4**: 4-7,

SCOTT RH (1966). Estimates of Hicksian IS and LM curves for the United States. *Journal of Finance* **21(3)**: 479-487

SIMON CP AND BLUME L (1994). *Mathematics for Economists.* Norton: New York.

SMITH GW (1991). Solution to a problem of stochastic process switching. *Econometrica* **59**: 237-239.

SMITH L (1995). Ensemble predictions and chaotic systems. Presented at the *6th International Meeting on Statistical Climatology*, University College, Galway.

SMITH LA (1994). Local optimal prediction: exploiting strangeness and the variation of sensitivity to initial condition. *Philosophical Transactions of the Royal Society of London A* **348**: 371-381.

SØRENSEN PH (1994). Option pricing in a Gaussian two-factor model of the term structure of interest rates. Research Paper, Copenhagen Business School, Denmark.

SPARROW C (1982). *The Lorenz Equations: Bifurcations, Chaos and Strange Attractors.* Springer-Verlag: New York.

STANTON R (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance* **52**. 1973-2002

STEELEY J (1989). Modelling the dynamics of the gilt edged term structure. Research Paper, Financial Options Research Centre, University of Warwick.

STEELEY JM (1990). Modelling the dynamics of the term structure of interest rates. *Economic and Social Review* **21**: 337-361.

STEVENSON A, MUSCATELLI V AND GREGORY M (1988). *Macroeconomic Theory and Stabilisation Policy.* Philip Allan: New York.

STIGLITZ JE (1992). Methodological issues and the new Keynesian economics. In: VERCELLI A AND DIMITRI (EDS). *Macroeconomics: A Survey of Research Strategies*: 38-86. Oxford University Press: Oxford.

STROUSTRUP B (1993). *The C++ Programming Language.* Addison-Wesley: Reading.

SUBBA RAO T (ED) (1993). *Developments in Time Series Analysis.* Chapman & Hall: London.

TATA F AND VASSILICOS JC (1991). Is there chaos in economic time series? A study of the stock exchange and the foreign exchange markets. Research Paper, LSE Financial Markets Group, London.

TAYLOR SJ (1994). Modelling stochastic volatility: a review and comparative study. *Mathematical Finance* **4**: 183-204.

TRIPPI RR AND TURBAN E (EDS) (1993). *Neural Networks in Finance and Investment.* Probus: Chicago.

VAN DIJK HK, MONFORT A AND BROWN BW (EDS) (1996). *Econometric Inference Using Simulation Techniques.* Wiley: Chichester.

VARIAN HR (1992). *Microeconomic Analysis, 3rd Ed.* Norton: New York.

VASICEK O (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics* **5**: 177-188.

VASSILICOS JC (1990). Are financial markets chaotic: a preliminary study of the foreign exchange market. Research Paper, Financial Markets Group, LSE.

VERHAEGEN M AND VAN DOOREN P (1986) Numerical aspects of different Kalman filter implementations. *IEEE Transactions on Automatic Control.* **31(10)**: 907-917

WAN EA (1993). Finite impulse response networks with applications in time series prediction. PhD Thesis, Stanford University.

WEBBER N (1992). Interest rate options and jump diffusion processes: a new approach to valuation. Research Paper, Warwick Business School.

WEDDEPOHL C (1995). A cautious price adjustment mechanism: chaotic behaviour. *Journal of Economic Behaviour and Organisation* **27**: 293-300.

WEI WWS (1990). *Time Series Analysis: Univariate and Multivariate Methods.* Addison-Wesley: Redwood City.

WEIGEND AS (1994). Paradigm change in prediction. *Philosophical Transactions of the Royal Society of London A* **348**: 405-420.

WIGGINS S (1990). *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* Springer-Verlag: New York.

WILLIAMSON J (ED) (1994). *Estimating Equilibrium Exchange Rates.* Institute for International Economics: Washington, DC.

WILMOTT P, DEWYNNE J AND HOWISON S (1994). *Option Pricing: Mathematical Models and Computation.* Oxford Financial Press: Oxford.

WINDER R (1993). *Developing C++ Software, 2nd Ed.* Wiley: Chichester.

WONNACOTT TH AND WONNACOTT RJ (1990). *Introductory Statistics for Business and Economics, 4th Ed.* Wiley: New York.

YAO Q AND TONG H (1994). On prediction and chaos in stochastic systems. *Philosophical Transactions of the Royal Society of London A* **348**: 357-369.

ZAPATERO F (1995). Equilibrium asset prices and exchange rates. *Journal of Economic Dynamics and Control* **19**: 787-811.

ZASLAVSKY GM, SAGDEEV RZ, USIKOV DA AND CHERNIKOV AA (1992). *Weak Chaos and Quasi-Regular Patterns.* Cambridge University Press: Cambridge.

ZHANG X, HUTCHINSON J (1994). Practical issues in nonlinear time series prediction. In: WEIGEND AS AND GERSHENFELD NA (EDS). *Time Series Prediction: Forecasting the Future and Understanding the Past.* Addison-Wesley: Reading.